

DISSERTATION THESIS
Charles University
Faculty of Science
Department of Biochemistry



Mgr. Kateřina Macháčková

Analogues of IGF-1 for the study of
interactions of the hormone with the
receptors for IGF-1 and insulin

Scientific Supervisor: RNDr. Jiří Jiráček, CSc.
Institute of Organic Chemistry and Biochemistry Czech Academy of Sciences

Prague, 2018

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 2. 06. 2018

Podpis

Acknowledgement

First of all, I would like to acknowledge my supervisor Jiří Jiráček. I'm grateful to him for accepting me as a PhD student under his wings and for teaching me with incredible patience a proper laboratory work, planning of experiments, a critical evaluation of obtained results as well as for showing me what the science is.

Secondly, I would like to thank all members, present or past, of our laboratory for sharing with me their experience and knowledge and also for their help with my experiments. Namely, I would like to thank Lenka Žáková, Jan Pícha and Květoslava Křížková.

Additionally, I would like to thank research-service groups of IOCB, namely, mass spectrometry group and NMR group.

Finally, I would like to express my deep gratitude to my family, mainly to my husband and my parents, for their psychological and financial support during my whole studies.

This work was supported by the Medical Research Council Grant MR/K000179/1, by the Medical Research Council Grant MR/R009066/1 and partly also by the Czech Science Foundation Grant 15-19018S. Institutional support was provided by the project RVO 61388963 (for the Institute of Organic Chemistry and Biochemistry).

1	Obsah	
2	Abstract.....	5
3	Abstrakt	6
4	List of Abbreviations	7
5	Introduction	8
5.1	The insulin/insulin-like growth factor (IGF) system	8
5.2	IGF-1	10
5.2.1	Structure.....	10
5.2.2	Binding to the receptors	11
5.2.3	Binding to IGF binding proteins	20
5.2.4	Biology of IGF-1	20
5.2.5	Diseases connected with IGF-1	23
5.3	Insulin-like growth factor binding proteins.....	24
5.4	The type 1 insulin-like growth factor receptor.....	26
5.4.1	Structure.....	26
5.4.2	Binding of hormones	28
5.4.3	Signalling cascade.....	32
5.4.4	IGF-1R internalization, recycling, and degradation	34
5.4.5	Nuclear IGF-1R	35
5.4.6	Hybrid receptors of IGF-1R.....	35
5.4.7	IGF-1R downregulation and diseases	36
5.5	Click chemistry	38
5.5.1	Cu ^I catalysed 1,3-dipolar cycloaddition of azides and alkynes	40
6	Research Aims.....	42
7	Publications	43
7.1	Insulin-like Growth Factor 1 Analogues Clicked in the C Domain: Chemical Synthesis and Biological Activities	44

7.2	Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides.....	90
7.3	Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain.....	113
7.4	Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation	136
8	Discussion.....	157
8.1	Fully synthetic IGF-1 analogues with modifications at the positions 36 and 37.....	157
8.1.1	Development of the new methodology for the total chemical synthesis of IGF-1 analogues and preparation of the analogues.....	157
8.1.2	Receptor binding and activation properties of fully synthetic IGF-1 analogues.	162
8.2	Development of a new methodology for the recombinant production of IGF-1, IGF-2 and their analogues in <i>E. coli</i>	164
8.3	Synthesis and characterization of IGF-1 analogues with mutations at the positions 45, 46 and 49.....	165
9	Conclusions	168
10	References.....	169

2 Abstract

Insulin/IGF system is a complex network of three similar hormones (insulin, IGF-1 and IGF-2) and their three similar receptors (IR-A, IR-B and IGF-1R, Figure 5.1.), which play important roles in maintaining basal energy homeostasis of the organism, in growth, development, life-span but also in development of diseases such as diabetes mellitus, cancer, acromegaly or Laron dwarfism. Despite structural similarities between family members, each member has its unique role in the system. Identification of structural determinants in insulin and IGFs that trigger their specific signalling pathways is important for rational drug design for safer treatment of diabetes or for more efficient combating of cancer or growth-related disorders. In this thesis, we focused on identification of such structural determinants in IGF-1. Comparison of our data with parallel studies with IGF-2 and insulin could give a more complex picture of the problem.

First of all, we developed necessary methodologies for the preparation of IGF-1 analogues. We developed a new methodology for the total chemical synthesis of IGF-1 analogues based on the solid-phase synthesis of fragments and their ligation by a Cu^I-catalyzed cycloaddition of azides and alkynes. In parallel, we developed a procedure for a recombinant production of IGF-1 and its analogues in *E. coli*.

Next, to gain an insight into the structural basis of IGF-1 binding specificity for IGF-1R, IR-A and IR-B, especially in comparison with insulin and IGF-2, we generated a series of mutants with specific amino acid substitutions at the positions 49, 45 and 46 of the B domain of IGF-1. In another study, we modified a pair of arginine residues at the positions 36 and 37 of the C domain of IGF-1. For all analogues we tested binding affinities of analogues for the selected receptors and abilities of analogues to activate these receptors.

Our data provided new insights into importance of the studied amino acids in IGF-1 for interaction with receptors for IGF-1 and insulin and may be useful for further rational engineering of new hormone analogues for potential medical applications.

3 Abstrakt

Insulin a hormony IGF jsou součástí komplexního systému tří navzájem si podobných hormonů (insulinu, IGF-1 a IGF-2) a jejich tří příbuzných receptorů (IR-A, IR-B a IGF-1R, obr. 5.1), který hraje důležitou roli v regulaci basální energetické rovnováhy organismu, v jeho růstu, vývoji, délce života ale i ve vývoji onemocnění jako je diabetes mellitus, rakovina, akromegalie nebo Laronův nanismus. Navzdory strukturním podobnostem má každý hormon svou jedinečnou roli v organismu. Identifikace specifických strukturních motivů v insulinu, IGF-1 a IGF-2, které spouštějí specifické signální dráhy, by byla přínosem pro racionální návrh léčiv, pro bezpečnější léčbu diabetu, či účinnější léčbu rakoviny nebo růstových onemocnění. V této práci jsme se zaměřili na identifikaci takových strukturních motivů u hormonu IGF-1. Srovnání dat se studii prováděných paralelně s IGF-2 a insulinem by mohlo poskytnout komplexnější pohled na danou problematiku.

Nejdříve jsme vyvinuly nezbytné metodologie pro přípravu analogů IGF-1. Vyvinuli jsme nový postup pro celkovou chemickou syntézu analogů IGF-1 založený na syntéze na pevné fázi fragmentů IGF-1 a jejich spojování pomocí cykloadice azidů a alkyňů katalyzované ionty Cu^I . Paralelně jsme vyvinuli postup pro rekombinantní produkci IGF-1 a jeho analogů v *E. coli*.

Abychom získali informace o důležitosti specifických strukturních motivů IGF-1 pro vazbu na IGF-1R, IR-A a IR-B zvláště ve srovnání s insulinem a IGF-2, připravili jsme série mutantů se substitucemi ve vybraných pozicích 49, 45 a 46 v doméně B a v pozicích 36 a 37 v doméně C IGF-1. U všech analogů jsme testovali jejich vazebné afinity vůči receptorů a rovněž jejich schopnosti tyto receptory aktivovat.

Naše data poskytla nové informace o důležitosti studovaných pozic v IGF-1 pro interakci s receptory IGF-1 a insulinu a mohou být užitečné pro další racionální modifikace těchto hormonů pro potenciální lékařské aplikace.

4 List of Abbreviations

ALS = acid labile subunit

GSH = reduced glutathione

GSSG = oxidized glutathione

IGF = insulin-like growth factor

IGFBPs = insulin-like growth factor binding proteins

IGF-1R = The type 1 insulin-like growth factor receptor

IGF-2R = The type 2 insulin-like growth factor receptor

IR = insulin receptor

IR-A = isoform A of the insulin receptor

IR-B = isoform B of the insulin receptor

IRSs = insulin receptor substrates

Nle = Norleucine

PI3K = phosphoinositol 3-kinase

PPP = cyclolignan picropodophyllin

RTK = receptor tyrosine kinase

siRNA = short-interfering RNA

5 Introduction

5.1 The insulin/insulin-like growth factor (IGF) system

Insulin/IGF system plays key developmental and metabolic roles at every stage of life. It is a complex system that controls and coordinates a number of biological effects such as cellular growth, tissue formation and remodelling, bone growth, brain development and basal energy metabolism. The cross-talk between components of insulin/IGFs system has a crucial role in proper coordination of these biological outcomes [1].

Insulin/IGF system consists of three receptors of the Insulin/IGF receptor tyrosine kinase subfamily (i.e. two isoforms of the insulin receptors, IR-A and IR-B, and the type 1 insulin-like growth factor receptor, IGF-1R), structurally different receptor for insulin-like growth factor 2 (IGF-2R), three hormones (IGF-1, IGF-2 and insulin), six circulating IGF binding proteins (IGFBP-1-6) and proteases that modulate IGF binding proteins availability [1, 2].

Although the insulin receptors (IR-A and IR-B) and IGF-1 receptor (IGF-1R) share similar structures and similar downstream cytoplasmic mediators it is known that IR-A and IR-B are involved mainly in a regulation of physiological processes such as glucose transport and metabolism, biosynthesis of glycogen and fat, whereas IGF-1R is mainly a regulator of cell growth, proliferation, and differentiation. Moreover, these receptors can form heterodimers called hybrid receptors [3].

All these receptors can be activated by the three hormones, IGF-1, IGF-2 and insulin (see Figure 5.1.) and each hormone has a different role in the organism [4]. It is supposed that IR-B is activated mainly by insulin, whereas IR-A mainly by insulin or IGF-2. IGF-1R is activated primarily by IGF-1 or IGF-2. Moreover IGF-2 can also bind to the IGF-2R, which is also known as mannose-6-phosphate receptor, and which functions are not fully clear [1].

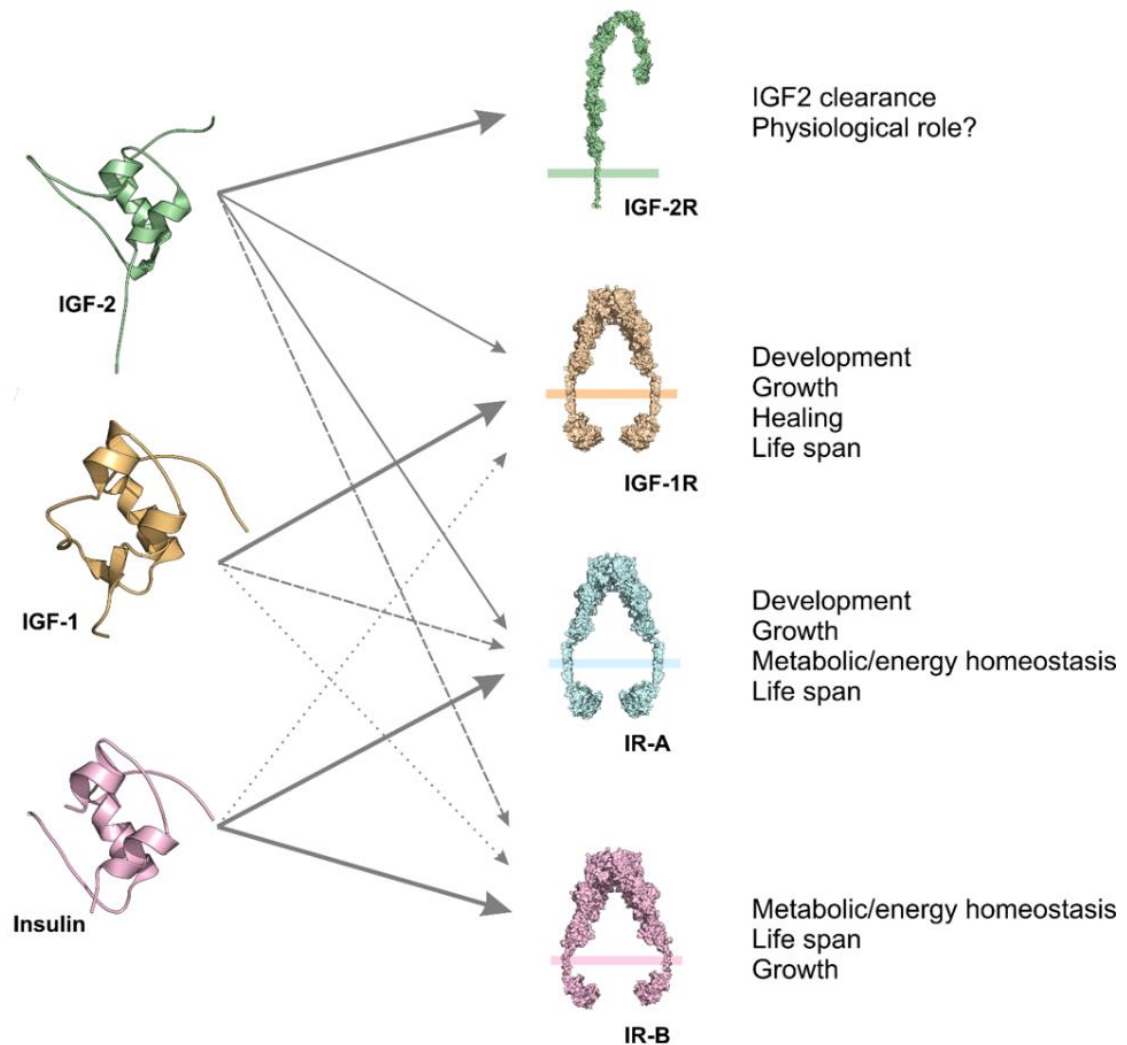


Figure 5.1. Insulin/IGFs system. Schematic and simplified representation of the binding of insulin, IGF-2 and IGF-1 to individual receptors and their predominant impacts on the organism. Solid and bold arrows show strong binding, thin and dashed arrows denote weaker binding. Typical affinities determined in the laboratory of Dr. Jiráček are shown in Table 5.1. (below). Figure was kindly prepared by Dr. Lenka Žáková.

Moreover, bioactivity of IGF-1 and IGF-2 in tissues is further regulated by circulating and local expression of these hormones, respective IGF-BPs and degrading proteases. In contrast, bioactivity of insulin is regulated by its production in β -cells of islets of Langerhans in pancreas in response to elevated glucose levels and other factors as well [5].

5.2 IGF-1

The insulin-like growth factor 1 (IGF-1) was discovered in the late 1950s as skeletal growth factor under control of the pituitary hormones produced in the liver that has impact on whole-body somatic growth and it was also known as somatomedin [6].

IGF-1 is a hormone, a member of insulin protein family as well as IGF-2 or insulin but also other human peptides as relaxins or INSL peptides. All these three hormones share similar secondary and tertiary structures and have similarities in primary structures as well, although their biological roles are different [6].

5.2.1 Structure

IGF-1 is 70 amino acid single chain polypeptide cross-linked by three disulphide bridges, that is divided into four domains: B, C, A and D (from N to C terminus, see Figure 5.2.). All four domains have about 60 % sequence identity with IGF-2, mostly in the B and A domains, which have also about 50 % sequence similarity with the B and A chains of insulin [6].



Figure 5.2. Primary sequences of human IGF-1, IGF-2 and insulin. Grey background highlights identical amino acids in all three hormones. Amino acids of B domains (B chain in insulin) is in a red box, C domain residues are in a green box, A domain (A chain in insulin) residues are in a blue box and D domain amino acids are in a purple box. [6]

The three dimensional structures of IGF-1 have been already solved by both NMR [7-10] and X-ray crystallography [11-14] methods (Figure 5.3.). The 3D-structure of IGF-1 consists of three highly conserved α -helices. Helix 1 (Gly8-Cys18) is in the B domain whereas helix 2 (Ile43-Cys47) and helix 3 (Leu54-Glu58) are both located in the A domain. The structure is further stabilized by three characteristic disulphide bonds (Cys6-Cys48, Cys18-Cys61 and Cys47-Cys52). The C and D domains appear highly flexible in the

solution structures determined by NMR [7-10]. Overall, the 3D structure is very similar to 3D structures of IGF-2 and insulin especially in the hydrophobic cores of hormones formed by helical parts (Figure 5.4.).

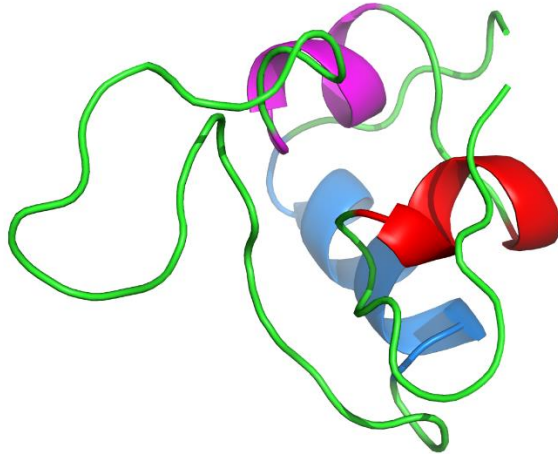


Figure 5.3. 3D-structure of IGF-1. Structure showing helix 1 (blue) in the B domain, helix 2 (pink) and helix 3 (red) in the A domain [9].

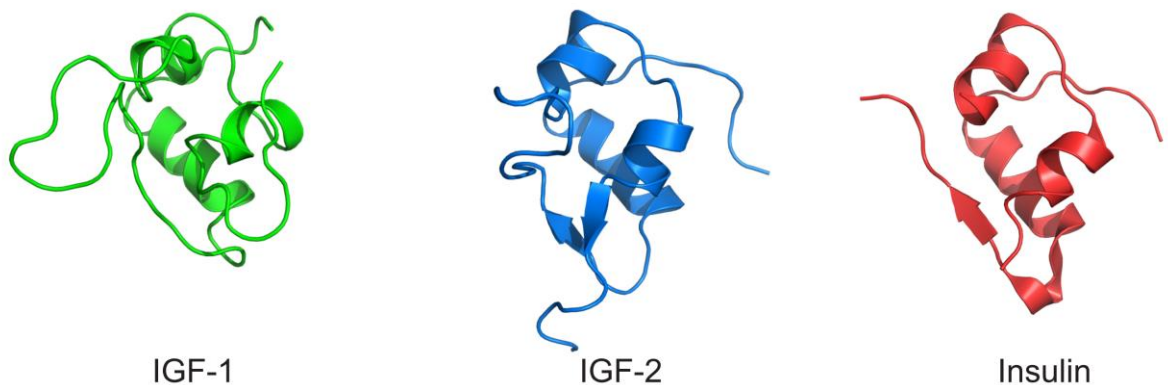


Figure 5.4 Comparison of 3D-structures of IGF-1, IGF-2 and insulin [9, 15, 16].

5.2.2 *Binding to the receptors*

IGF-1 binds to the insulin/IGF subfamily of cell surface tyrosine kinase receptors. IGF-1 binds IR-A (without 12 amino acids at the C-terminus of the α -domain) and especially IR-B (with 12 amino acids at the C-terminus of the α -domain) with a low affinity. Typical

binding affinities of IGF-1 but also of IGF-2 and human insulin determined in the laboratory of Dr. Jiráček are shown in Table 5.1.

Table 5.1. Binding affinities of insulin, IGF-1 and IGF-2. Typical values of binding affinities (K_d values) of insulin, IGF-1 and IGF-2 for human IR-A (in membranes of human IM-9 lymphocytes), human IR-B or human IGF-1R (in membranes of mouse fibroblasts) measured in Dr. Jiráček's laboratory using ^{125}I -labeled insulin or ^{125}I -labeled IGF-1 as tracers. Adapted from [17].

Hormone/Receptor	IR-A (nM)	IR-B (nM)	IGF-1R (nM)
Human IGF-1	24	225	0.2-0.3
Human IGF-2	2.9-3.0	35	2.3
Human insulin	0.2-0.5	0.3-0.7	290

IGF-1 binds also with relatively high affinities both hybrid receptors, IGF-1R/IR-A (EC_{50} about 0.3 nM) or IGF-1R/IR-B (EC_{50} about 2.5 nM) [18].

Different binding affinities of IGF-1 for IR and IGF-1R receptors indicate different binding interactions with each receptor, however details of these interactions are still only partly understood. Interaction of IGF-1 especially with IGF-1R and IR-A receptors is extensively investigated through the preparation and functional analysis of IGF-1 analogues. Table 5.2. shows a list of IGF-1 analogues and their binding properties, which have been already prepared. There is a variability of the absolute values of binding affinities of analogues from different studies due to different techniques used by different groups to determine binding affinities. However, the relative binding affinities (related to wild type IGF-1), are relatively consistent between individual studies. Moreover, some early studies with IR were performed with different tissues often expressing a mixture of both IR isoforms.

Chimeras combining domains of insulin and IGF-1 or domains of IGF-1 and IGF-2 were frequently used to identify regions important for differential binding to the receptors. Likewise, point mutations, deletions or multiple mutations were frequently used for mapping binding epitopes in insulin, IGF-1 and IGF-2. At least 24 of the 70 IGF-1 amino acids have been individually mutated and tested for their receptors binding properties.

Experiments involving chimeras showed, that the C domain plays a role in determining the IGF-1R binding preference for IGF-1. Substitution of the IGF-1 C domain for the IGF-2 C domain lead to the IGF-1R binding affinity similar to that of native IGF-2

and reverse effect was observed for IGF-2 containing the IGF-1 C domain [19]. Likewise, it was found that binding of IGF-1 analogue is disrupted when using IGF-1R chimera with IR CR domain suggesting that the IGF-1 C domain interacts predominantly with IGF-1R CR domain [20]. Mutations of Tyr31, Arg36 and Arg37 revealed their importance for IGF-1R binding. Substitution of Arg36 and Arg37 with Ala leads to a decrease of binding to IGF-1R [20, 21] 5- to 15- fold, depending on the assay methodology. Mutation of Tyr31 to Ala caused a 6-fold decrease of binding to IGF-1R [22].

Ala mutations of IGF-1 residues Tyr31, Arg36 and Arg37 cause increase in IR binding [20-22], which can support the assumption that the C domain of IGF-1 plays a rather minor role in IR binding. Another experiments leading to this conclusion involved replacement of the IGF-1 C domain with a 4-glycine bridge and caused an increase binding to IR [23], a single chain insulin with the IGF-1 C domain had IR affinity similar to native human insulin [24], and double chain insulin with the IGF-1 C domain on the C-terminus of the B chain had a lower IR affinity than native insulin [25]. In contrast, it seems that IGF-2 C domain has more important role in IR binding [19].

Although the IGF-1 D domain is less significant in IGF-1R binding, a study mutating the positively charged Lys65 and Lys68 residues in the IGF-1 D domain for Ala showed their importance in IR binding properties of IGF-1. Their mutation to Ala caused an increase in IR binding potency relative to IGF-1 indicating that they can represent an obstacle in binding to the IR [20].

Both A and B domains, in general, have crucial roles in binding properties of IGF-1 to both IR isoforms and to IGF-1R receptor. However, it was shown that the N-terminal tail of IGF-1 B domain is not crucial for IGF-1R binding. Sequential deletions of the N-terminus of IGF-1 revealed that removal of first 4 residues do not significantly disrupt IGF-1R binding [26]. Mutagenesis of Glu3 and Thr4 supported this conclusion [27-30]. However, Thr4 has influence on IR binding since a mutation of this residue to His, which is the corresponding residue in insulin, led to an increase in affinity for the IR-A and IR-B [28].

Nevertheless a few residues, namely Ala8, Val11, Phe23 and Tyr24, of the B domain were shown as crucial for binding of IGF-1 to both IR and IGF-1R receptors as well [31].

An additional interesting position in IGF-1 is Glu9. Analogue of IGF-1 with the substitution of negatively charged Glu9 to positively charged Lys9 binds IGF-1R with 25 % of native IGF-1 affinity and Glu9Gln IGF-1 binds IGF-1R 75 % of native IGF-1 [32]. Interestingly, the mutation of corresponding HisB10 residue in insulin to Lys also reduces

IR binding of the analogue [1]. However, mutation HisB10Asp of insulin results in insulin analogue with very high potency for both receptors [33]. This suggests the preference of both receptors, IGF-1R and IR-A, for a negative residue or for the lack of a positive charge at this position at the beginning of the B chain helix.

Moreover, it was determined that the substitution of the first 16 N-terminal IGF-1 amino acids with the first 17 N-terminal amino acids of insulin B-chain, which includes 10 sequence changes, has almost no effect on IGF-1R binding [27]. These data indicate that the IGF-1R binding surface is similar within the B domains of insulin and IGF-1.

IGF-1 A domain interactions were not investigated as intensively as the B domain interactions. Nevertheless, residues Val44 [34], Glu58 [35], Met59 [28] and Tyr60 [22] were identified as crucial for IGF-1R binding. Val44Met IGF-1 has been found in a patient with severe mental and growth retardation, which was explained by the impaired IGF-1R binding [36]. On the other hand, A domain residues that were not shown to be important for IGF-1R binding include Phe49 [29], Asp53 [35], Arg50, Arg55 and Arg56 [21]. Residues Ala54 and Leu55 in IGF-2 that are equivalent to Arg55 and Arg56 in IGF-1 also do not seem to be crucial for IGF-1R binding [37].

Exchanging residues 42–56 in IGF-1 for the first 15 N-terminal residues of the insulin A chain, together with a substitution of Thr41 for Ile in the C domain, led to a 7-fold increase in IR binding thus indicating an important role of these residues in IR binding. On the other hand, this IGF-1 chimera maintained IGF-1R binding affinity similar to native IGF-1, suggesting different binding epitopes in IGF-1R and IR responsible for binding of these parts of the hormones. Interestingly, the replacement of Phe49, Arg50 and Ser51 or the replacement of Arg55 and Arg56 in IGF-1 with the equivalent residues in insulin do not change the binding affinity to IR. However, it was found the simultaneous replacement of all these residues led to an increase of IR binding [38].

Table 5.2. IGF-1R, IR-A, IR-B and IGF-2R binding affinities of IGF-1 analogues. Binding affinities are relative to IGF-1. B domain is in blue, C domain is in green, A domain is in black and D domain is in red.

N.B.= no binding detected, *13 amino acid N-terminal extension (M F P A M P L S S L F V N)

1	10	20	30	40	50	60	70	IR-A	IR-B	IGF-1R	IGF-2R	Reference
GPETLCGAE	LVDALQFVCG	DRGFYFNKPT	GYGSSRRAP	QTGIVDECCF	RSCDLRRLEM	YCAPLKPAKSA		28.9		0.07		[20]
-----	-----	-----	-----AA-----	-----	-----	-----	-----	5.77		0.10		[20]
-----	-----	-----	-----	-----	S-----YE-----	-----	-----	0.79		0.83		[20]
-----	-----	-A-----	-----	-----	-----	-----	-----			0.32		[21]
-----	-----	-----	-----AA-----	-----	-----	-----	-----			0.18		[21]
-----	-----	-----	-----	-----	A-----	-----	-----			0.64		[21]
-----	-----	-----	-----	-----	-----A-----	-----	-----			0.91		[21]
-----	-----	-----	-----	-----	-----A-----	-----	-----			0.51		[21]
-----	-----	-----	-----	-----	-----	-----	-----			NB		[39]
-----	-----	-----	GGGG	-----	-----	-----	-----			0.01		[39]
-----	-----	-----	SR VSRRS	R-----	-----	-----	-----	1.88	2.04	0.25		[19]
-----	-----	-----	-----	-----	-----	-----T--PAKSE	-----	1.44	1.24	1.14		[19]
-----	-----	-----	SR VSRRS	R-----	-----	-----T--PAKSE	-----	6.15	8.32	1.11		[19]
-----	-----	-----	GGGG	-----	-----	-----	-----	1.75		0.03	0.44	[23]
-----	-----	-----	-----	-----	-----	-----	-----	2		0.77	1	[23]
-----	-----	-----	GGGG	-----	-----	-----	-----	4.67		0.02	0.40	[23]
FVNQHLCGSH	LVEALYLVCG	ER-----	-----	-----GIVEQCCT	SICSLYQLEN	YCN-----	-----	46		0.17		[25]
-----	-----	-----	-----	-----M-----	-----	-----	-----			0.0001		[34]
-----H-----	-----Y-----	-----	-----	-----T-----	-----I-----	-----	-----	41.4				[40]
-----H-----	-----	-----	-----	-----	-----	-----	-----	4.07				[40]
-----	-----Y-----	-----	-----	-----	-----	-----	-----	3.82				[40]
-----	-----	-----	-----	-----T-----	-----	-----	-----	1.04				[40]

1	10	20	30	40	50	60	70	IR-A	IR-B	IGF-1R	IGF-2R	Reference
GPETLCGAE	LVDALQFVCG	DRGFYFNKPT	GYGSSRRAP	QTGIVDECCF	RSCDLRRLEM	YCAPLKPAKSA		0.61				[40]
-----	-----	-----	-----	-----	-I-----	-----						
FVNQHLCGSH	LVEALYL---	-----	-----	-----	-----	-----		4		0.47	<0.06	[27]
--QA-----	-----	-----	-----	-----	-----	-----		2		1.05	0.3	[27]
-----	-----YL---	-----	-----	-----	-----	-----		9.33		1.24	3.75	[27]
--QA-----	-----YL---	-----	-----	-----	-----	-----		14		1.11	2.72	[27]
--H-----	-----	-----	-----	-----	-----	-----				1.88		[32]
-----Q---	-----	-----	-----	-----	-----	-----				0.72		[32]
-----K---	-----	-----	-----	-----	-----	-----				0.21		[32]
-----	-----S---	-----	-----	-----	-----	-----				1.29		[32]
-----Q---	-----S---	-----	-----	-----	-----	-----				0.59		[32]
-----K---	-----S---	-----	-----	-----	-----	-----				0.32		[32]
-----	-----	-----	-----	-----	-----	-----				1		[26]
-----	-----	-----	-----	-----	-----	-----				1		[26]
-----	-----	-----	-----	-----	-----	-----				1.5		[26]
-----	-----	-----	-----	-----	-----	-----				0.5		[26]
-----	-----	-----	-----	-----	-----	-----				<0.01		[26]
-----	-----	-----	-----	-----	-----	-----		1.4		2.5	<0.001	[41]
--R-----	-----	-----	-----	-----	-----	-----		0.65		0.8	<0.001	[41]
--G-----	-----	-----	-----	-----	-----	-----				1.28		[42]
--R-----	-----	-----	-----	-----	-----	-----				1.25		[42]
-----K---	-----	-----	-----	-----	-----	-----				0.17		[42]
-----A---	-----	-----	-----	-----	-----	-----				0.34		[35]
-----	-----A---	-----	-----	-----	-----	-----				0.29		[35]
-----	-----A---	-----	-----	-----	-----	-----				0.50		[35]
-----	-----	-----	-----	-----	-----A---	-----				0.73		[35]
-----	-----	-----	-----	-----	-----A---	-----				0.24		[35]

1	10	20	30	40	50	60	70	IR-A	IR-B	IGF-1R	IGF-2R	Reference
GPETLCGAE	LVDALQFVCG	DRGFYFNKPT	GYGSSRRAP	QTGIVDECCF	RSCDLRRLEM	YCAPLKPAKSA				0.15		[35]
-----	-----	-----	-----	-----	-----	-----A-----	-----			<0.03		[35]
-----A-----	-----A-----	-----	-----	-----	-----	-----	-----			<0.01		[35]
-----A-----	-----	-----	-----	-----	-----	-----A-----	-----			1.76		[43]
-----R-----	-----	-----	-----	-----	-----	-----	-----			0.83		[43]
-----G-----	-----	-----	-----	-----	-----	-----	-----			1.32		[43]
-----	-----	-----	-----	-----	-----	-----	-----			0.05		[43]
-----	-----	-----L-----	-----	-----	-----	-----	-----			0.03		[43]
-----	-----	-----L-----	-----	-----	-----	-----	-----			0.08		[43]
-----	-----	-----	-----A-----	-----	-----	-----	-----			0.09		[43]
-----	-----	-----	-----A-----	-----	-----	-----	-----			0.005		[43]
-----	-----	-----	-----	-----	-----	L-----	-----			NB		[43]
-----	-----	-----	-----A-----	-----	-----	L-----	-----			<0.085		[44]
-----	-----	-----G-----	-----	-----	-----	-----	-----			0.04	<0.11	[28]
-----L-----	-----	-----	-----	-----	-----	-----	-----			0.71	0.79	[28]
-----	-----S-----	-----	-----	-----	-----	-----	-----			0.12	0.20	[28]
-----	-----S-----	-----	-----	-----	-----	-----L-----	-----			0.50	0.48	[28]
-----	-----	-----	-----	-----	-----	-----L-----	-----			6.7	3.5	[28]
-----H-----	-----	-----	-----	-----	-----	-----	-----			0.06	<0.11	[28]
-----	-----	-----	-----	-----	-----	-----F-----	-----			0.20	0.05	0.78
-----	-----	-----L-----	-----	-----	-----	-----	-----			1.58	0.17	0.88
-----	-----	-----	-----A-----	-----	-----	-----	-----			<0.57	0.05	<0.7
-----	-----	-----	-----A-----	-----	-----	L-----	-----			<0.57	0.002	<0.7
-----	-----	-----	-----A-----	-----	-----	L-----	-----			<0.57	0.004	2.33
-----	-----	-----L-----	-----	-----	-----	-----	-----			<0.57	0.001	<0.7
-----	-----	-----L-----	-----	-----	-----	L-----	-----			<0.57	0.001	<0.7
-----	-----	-----L-----	-----A-----	-----	-----	L-----	-----			<0.57	0.001	<0.7

1	10	20	30	40	50	60	70	IR-A	IR-B	IGF-1R	IGF-2R	Reference
GPETLCGAE	LVDALQFVCG	DRGFYFNKPT	GYGSSRRAP	QTGIVDECCF	RSCDLRRLEM	YCAPLKPAKSA		0.11		0.30		[45]
-----	-T-----	-----	-----	-----	-----	-----		0.26		1.57		[45]
-----	-I-----	-----	-----	-----	-----	-----		0.38		0.92		[45]
-----	-A-----	-----	-----	-----	-----	F-----				0.40		[46]
-----	-A-----	-----	-----	-----	-----	-----				0.27		[46]
-----	-A-----	-----	-----	-----	-----	-----				0.51		[46]
-----	-E-----	-----	-----	-----	-----	-----				0.69		[46]
-----	-A-----	-----	-----	-----	-----	-----				0.03		[46]
-----	-----	-S-----	-----	-----	-----	-----		0.45		0.06	0.50	[47]
-----	-----	-L-----	-----	-----	-----	-----		0.10		0.03	0.33	[47]
-----	-----	-FFY-----	-----	-----	-----	-----		1.87		0.73	3	[47]
-----	-----	-L-----	-----	-----	-----	-----		0.22		0.01	1	[47]
-----	-----	-----	-----	-----	-A-----	-----				1		[29]
-A-----	-----	-----	-----	-----	-A-----	-----				1		[29]
-----	-----	-----	-----	-IGIVEQCCT	SICSLYQ---	-----		7		1.53	<0.08	[38]
-----	-----	-----	-----	-----	-----	-YQ-----		0.82		1.04	6.67	[38]
-----	-----	-----	-----	-----	-----	-T SI-----		1.27		0.70	<0.04	[38]
--R-----	-----	-----	-----	-----	-----	-----				0.63		[30]
--G-----	-----	-----	-----	-----	-----	-----				0.52		[30]
*-------	-----	-----	-----	-----	-----	-----				1		[48]
*--G-----	-----	-----	-----	-----	-----	-----				0.49		[48]
*--R-----	-----	-----	-----	-----	-----	-----				0.23		[48]
*-------	-----	-----	-----	-----	-----	-----				0.70		[42]
*--G-----	-----	-----	-----	-----	-----	-----				0.68		[42]
*--R-----	-----	-----	-----	-----	-----	-----				0.40		[42]
*--K-----	-----	-----	-----	-----	-----	-----				0.22		[42]

1	10	20	30	40	50	60	70	IR-A	IR-B	IGF-1R	IGF-2R	Reference
GPETLCGAE	LVDALQFVCG	DRGFYFNKPT	GYGSSRRAP	QTGIVDECCF	RSCDLRRLEM	YCAPLKPAKSA		1.2		0.31		[49]
-----	-----	-----	-----	-----H-----	-----	-----		1.2		0.60		[49]
-----	-----	-----	-----	-----N-----	-----	-----		1.3		0.23		[49]
-----	-----	-----	-----	-----A-----	-----	-----		3.5		1.33		[49]
-----	-----	-----	-----	-----H-----	-----	-----		1.3		1.33		[49]
-----	-----	-----	-----	-----Q-----	-----	-----		1.7		0.92		[49]
-----	-----	-----	-----	-----A-----	-----	-----		1.4		0.49		[49]
-----	-----	-----	-----	-----NQ-----	-----	-----		3.5		0.65		[49]
-----	-----	-----	-----	-----H-----	-----	-----		7		0.55		[49]
-----	-----	-----	-----	-----H-H-----	-----	-----		3.2		0.53		[49]
-----	-----	-----	-----	-----Q-H-----	-----	-----		4.3		0.37		[49]
-----	-----	-----	-----	-----NQ-H-----	-----	-----						

5.2.3 Binding to IGF binding proteins

IGF-1 binds to insulin-like growth factor binding proteins with high affinities (see table 5.3.), which are different for each of proteins [32].

Table 5.3. Binding affinities of IGF-1 for insulin-like growth factor binding proteins. From Ref. [32].

Protein	ED50 of IGF-1 (nM)
Human IGFBP-1	8.1 ± 0.7
Rat IGFBP-2	0.18 ± 0.02
Human IGFBP-3	0.08 ± 0.02
Human IGFBP-4	0.92 ± 0.02
Human IGFBP-5	0.65 ± 0.03
Human IGFBP-6	2.1 ± 0.1

In general, specific IGF-1 mutants have different decreases in IGFBP binding affinities compared to nature IGF-1, depending on the IGFBP investigated, which suggests that the IGFBP-binding epitopes in IGF-1 can differ for each IGFBP [50-52]. Since the most frequent serum IGFBP is IGFBP-3, it is also the most studied for IGF-1 binding. Truncated des-(1-3)-IGF-1 binds to IGFBP-3 with several times lower affinity than natural IGF-1 and has also reduced binding affinity to other IGFBPs [53]. Other data measured on IGF-1 Glu3 mutants showed that the Glu3 is an important determinant of the des-(1-3)-IGF-1 binding properties [29, 41].

Alanine scanning mutagenesis of the entire IGF-1 molecule identified Val11, Arg36, Pro39 and Pro63 like other determinants important for IGFBP-3 binding. The same residues were identified to be important also for IGFBP-1 binding [51].

Other IGF-1 residues found to be important for IGFBP binding are Glu15, Phe16, Phe48, Arg49, and Ser50 [27, 32, 52, 54].

5.2.4 Biology of IGF-1

IGF-1 is a hormone, with biological effects typical for a growth factor, which acts either in autocrine or paracrine fashion. It is synthesized in multiple tissues including liver, skeletal muscle, bone and cartilage. Since the majority of IGF-1 is synthesized by the liver, this organ is the most important regulator of blood concentration of IGF-1 and about 80 % of the total serum IGF-1 comes from the liver. Circulating IGF-1 acts by a paracrine

mechanism. The remaining IGF-1 is synthesized in the periphery, usually by connective tissue cell types. IGF-1 that is synthesized in the periphery locally regulates cell growth by an autocrine mechanism (it can bind to receptors that are present on the connective tissue cells themselves) and by paracrine mechanism as well (can bind to receptors on adjacent cell types, often epithelial cell types, that do not actually synthesize IGF-1) [55]. Moreover, the mesenchymal cells surrounding the tumours are also an important paracrine source of IGF-1 [56].

The synthesis of IGF-1 by cells in the connective tissue is under control of GH (the human pituitary growth hormone, also known as somatotropin). Secretion of growth hormone (GH) in the pituitary is regulated by the neurosecretory nuclei of the hypothalamus [57].

Serum IGF-1 circulates in relatively high concentrations (150–400 ng per ml) in plasma, mostly as the IGFBP-bound form. The free active form represents only a small part (less than 1 %) [58]. Serum levels of IGF-1 increase slowly from birth to a pubertal peak and then decrease with age. IGF-1, in contrast to IGF-2, is expressed at a low level embryonically and thus seems to be more important for postnatal growth and development, however it is still needed for correct embryonic development [59]. For example, foetal brain expression of IGF-1 is suggested to be crucial for determining brain growth and head size [60].

In mice, germline deletion of *Igf1* gene results in a 50 % reduction in birth weight and a 70 % reduction in final adult size [61], whereas selective deletion of IGF-1 synthesis in the liver (without effect on peripheral synthesis), results in less than a 10 % reduction in adult size [55]. These data support the theory that the peripheral synthesis of IGF-1, stimulated by GH, is an important determinant of somatic growth.

Hepatic synthesis of IGF-1 is regulated by several factors, including the GH hormone, but the ability of GH to stimulate synthesis of IGF-1 is strongly influenced by nutritional status [62]. In addition, increasing blood IGF-1 concentrations suppress GH synthesis in the pituitary gland (a process termed negative-feedback regulation) [63].

Other hormones, such as thyroxine, cortisol, estradiol and testosterone participate with GH in regulation of hepatic IGF-1 synthesis. Thyroxine increases sensitivity to GH, and thus IGF-1 synthesis. Whereas estradiol, as well as cortisol, inhibit IGF-1 synthesis in the liver. High cortisol concentrations can lead to growth attenuation by this mechanism. Testosterone enhances hepatic IGF-1 synthesis, moreover alters the sensitivity of the

pituitary gland to negative-feedback regulation of GH secretion, and thus causes an increase in both GH and IGF-1 [64].

Increasing of IGF-1 synthesis is also connected with injury. A wave of IGF-1 synthesis that stimulates reparative cell types to replicate was observed following tissue or cellular injury of blood vessels [65], skeletal muscle [66], cartilage [67] and of the brain [68]. Additionally, local synthesis of IGF-1 can be elicited by other growth factors involving in the repair process, such as platelet derived growth factor (PDGF), fibroblast growth factor (FGF) and epidermal growth factor (EGF) [69].

IGF-1 as a growth hormone plays a significant role also in protein metabolism. It was demonstrated that IGF-1 enhances protein synthesis and slows down protein degradation not only in muscle and skeletal tissues. In general, it mediates the anabolic effects [70-72].

Although IGF-1 is a growth hormone, it plays also another roles in organism than only growth and development [2]. Structural similarities between IGF-1, IGF-2 and insulin suggest that both peptides may arise from a common evolutionary precursor, which was used by organisms to link nutrient intake and growth. The IGF-1/IGF-2/insulin precursor could be synthesised after the food intake to stimulate cells to use the ingested nutrient for protein synthesis and tissue growth. IGF-1 has retained some insulin-like properties such as stimulation of glucose transport into skeletal muscle cells and its synthesis is linked to nutrient intake [73].

Several studies suggested that IGF-1 can enhance insulin action. For example glucose disposal is partially dependent on circulating IGF-1 concentrations [74] and deletion of the IGF-1R in skeletal muscle of mice results in glucose intolerance connected with impaired insulin action [75]. The mechanisms of these IGF-1 actions is proposed to be connected with either cross-talking between the insulin and the IGF-1 signal transduction pathways or with inhibiting the secretion of GH, which can act as an insulin antagonist [76].

IGF-1/IGF-1R has also unknown role in the brain connected not only with controlling synthesis of GH by the neuroendocrine brain. IGF-1 is also synthesized in the brain, within neurons and glia. Partial inactivation of IGF-1R in the embryonic brain selectively inhibited GH and IGF-1 pathways after birth, caused growth retardation, smaller adult size, and metabolic alterations, and led to delayed mortality and longer mean lifespan. It suggests that the development of this hormone axis is under control of the central nervous system (CNS) [77]. The brain IGF-1 action is mostly connected with lifespan [78]. Neurons in the central nervous system, by sensing the circulating levels of ligand, may have a central function in

regulating the ageing of all tissues. This possibility was investigated using the Cre-lox approach to produce brain-specific IGF-1R knockout mice. Homozygous mice for this mutation are microcephalic, sterile, and have a complex neuroendocrine dysfunction, but heterozygous mice are healthy and have an increase lifespan [77].

5.2.5 *Diseases connected with IGF-1*

Defects of the IGF-1 gene are rather rare in humans. Only a few cases (three homozygous and two families with heterozygous mutations) have been described. Patients have a variable degree of intrauterine and postnatal growth retardation, microcephaly, deafness and developmental delay [79]. However a relatively high concentration of serum IGF-1 is connected with an increased risk of developing cancer [80].

On the other hand, defects in GH action are not as rare as IGF-1 and lead to diseases such as dwarfism or acromegaly. Children with attenuation of GH action often also have IGF-1 deficiency that could be remedied by IGF-1 replacement. These treatments should be beneficial especially in cases of GH receptor defects that resulted in the resistance to GH. Several clinical studies used recombinant human IGF-1 (sold as Mecermin) for treatment children with IGF-1 deficiency [81-83]. Although the long term effect of this therapy did not allowed children to achieve a normal range of adult heights, their adult heights is significantly greater than in the absence of the therapy. Conceivably, further improvements such as initiating treatment earlier, or using different dosing regimens, could lead to better results [83]. IGF-1 has also been co-administered with IGFBP-3. Treatment of children with this complex results in more stable serum IGF-1 levels with growth rates in the first year of therapy equivalent to those obtained using IGF-1 alone [81].

Several catabolic states were described to lead to a relative resistance to GH as a result of increased concentrations of cytokines, which thus inhibit IGF-1 synthesis and block its actions in tissues [84]. These patients as well as patients with burns treated by IGF-1 respond to IGF-1 with increases in protein synthesis and an enhancement of the whole body protein anabolism [85, 86]. Likewise, experiments performed on rats showed that treatment by injection of biotinylated nanofibers delivering IGF-1, which provided sustained IGF-1 delivery for 28 days into rat myocardium, could be beneficial after myocardial infarction [87].

Two studies demonstrated that the reduction in bone mineral density can be improved by administered IGF-1, as a potent growth factor for osteoblasts [88, 89]. In addition, studies on animal models show that IGF-1 could be involved in stimulating atherogenesis [90, 91].

Administering IGF-1 to patients with severe insulin resistance, diabetes mellitus type 1 or type 2 lead to enhanced insulin sensitivity and improvement of post-prandial glucose usage [74, 92-94].

In addition, IGF-1 has a role in longevity connected with its action in brain. It was reported that heterozygous knockout mice (*Igf1r^{+/-}*) (heterozygous because null mutants are not viable) live on average 26 % longer than their wild-type littermates. For females, the effect was even more significant than for males. Long-lived *Igf1r^{+/-}* mice do not develop dwarfism, and their nutrient uptake, energy metabolism, physical activity, fertility and reproduction are unaffected. Moreover they display greater resistance to oxidative stress, a known determinant of ageing [95]. The same was observed in nematodes [96] (*Caenorhabditis elegans* [97]) and in the insect *Drosophila melanogaster* [98]. The presence of low levels of GH/IGF-1 in most of the long-lived rodent models delays puberty and reproduction. This observation led to an idea of shifting a 'biological timer' that will also delay aging. Thus, the longer lifespan of these models may be due to the late arrival of age-related diseases, despite many other problems that they face early in life [99].

Finally, IGF-1 was demonstrated to play a role in Alzheimer's disease [100] and in Huntington's disease [101].

5.3 Insulin-like growth factor binding proteins

Insulin-like growth factor binding proteins family consists of six protein members (IGFBP-1 to -6), which bind IGF-1 and IGF-2 with high affinity. They are well known for modulating bioavailability of IGF-1 and IGF-2, whereas their bioavailability is known to be modulated by proteolysis [102].

In serum, IGF-1 and IGF-2 are predominantly (about 90 %) found in a ternary complex consisting of IGF-1 or IGF-2, IGFBP-3 and acid labile subunit (ALS). Another ternary complex of IGF:IGFBP-5:ALS is also formed in a lower amount (about 10 %). ALS is 85 kDa glycosylated protein and belongs to the leucine-rich repeats (LRR) family of proteins, which main role is to increase the half-life of IGFs in the circulation. The downregulation of ALS leads to low serum IGFBP-3 and IGF-1, which causes growth retardation [103]. Surprisingly, ALS binds IGFBPs only if they are in complex with IGF [104, 105].

Both of these ALS-containing complexes work as a reservoir of IGFs trapped in circulation because they are not able to cross the vascular epithelial layer. Moreover these complexes have half-life about 16 hours, in contrast to less than 15 minutes of free IGF-1 [106].

The rest of IGFBPs (IGFBP-1, -2, -4, -6) are also present in serum, however in small amounts (the most frequent is IGFBP-2) and form smaller binary IGF:IGFBPs complexes, which can cross the vascular epithelial layer, and thus deliver IGF-1 to target tissues. Binding of these IGFBPs to IGFs also considerably slows down serum IGFs clearance [107, 108]. Likewise, locally expressed IGFBPs also modulate IGFs actions by blocking their binding to the IGF-1R [102].

In addition, almost all IGFBPs have been shown to have also IGF-independent role in organism, however their IGF-independent actions are much less investigated than the IGF-dependent actions [109, 110].

The mature IGFBPs contain 216–289 amino acids that are organized into three domains of approximately equal size. There are more conserved cysteine-rich N- and C-domains, which are linked together via L-domain, which is unique to each IGFBP. IGFBPs 1–5 have 18 conserved cysteines, whereas IGFBP-6 has 16 cysteines resulting in its different fold from other IGFBPs. All six IGFBPs have the same three conserved disulphide bonds within the C-domains [102, 111, 112].

Both conserved regions, the N- and C-domains, are binding IGFs and the IGF binding determinants are similar in all six IGFBPs [111, 113-116]. The binding site for ALS is located in the C-domain [117].

The L-domains of IGFBPs are not conserved and are not directly involved in binding of IGFs. They contain mainly sites of post-translational modifications such as glycosylation, phosphorylation and proteolysis. Glycosylation of IGFBPs does not change IGF binding affinities, but influences other IGFBPs properties such as stability, circulating half-life, and susceptibility to proteolysis and cell association [118-120]. Phosphorylation has been also shown to influence IGFBPs properties such as susceptibility to proteolysis [121, 122] or cell surface binding [123]. Moreover, in contrast to glycosylation, phosphorylation was demonstrated to influence IGF binding properties [124].

Proteolysis of IGFBPs predominantly occurs in the L-domain and provides a mechanism for regulation of circulating IGFBPs amounts. Moreover, some of the resulting fragments of the proteolysis have been reported to retain biological activity [125, 126]. A

number of proteases cleaving IGF-BPs was identified, such as plasmin, pregnancy-associated plasma protein A (PAPP-A), pepsin, cathepsins or matrix metalloproteinases [102, 112, 127-130].

5.4 The type 1 insulin-like growth factor receptor

The type 1 insulin-like growth factor receptor (IGF-1R) is a transmembrane tyrosine kinase that belongs to insulin/IGF-1 receptor subfamily of receptor tyrosine kinases (RTK) [131, 132]. All RTKs share a similar molecular architecture, with large extracellular glycosylated ligand-binding domain, a single transmembrane helix, and intracellular region that contains a juxtamembrane regulatory region, protein tyrosine kinase domain with an additional C-terminal region. They are one of a major cell surface receptor family, containing about 60 members, subdivided into at least 13 receptor subfamilies. The major common structural feature of these receptors is their dimeric character or dimerization upon activation by a ligand [131, 133].

IGF-1R is similar to IR-A (insulin receptor isoform A, lacking 12 amino acids at the C-terminus of the α -subunit coded by exon 11) and to IR-B (insulin receptor isoform B, having 12 amino acids at the C-terminus of the α -subunit coded by exon 11). The IGF-1R and the IRs have high sequence homology (84 % in the tyrosine kinase domain, about 55 % in the ligand binding part, and above 50 % in overall amino acid sequence) [134].

IGF-1R binds IGF-1 with high affinity and IGF-2 with lower affinity as well but insulin with a low affinity (Table 5.1.) [17].

5.4.1 Structure

The human IGF-1R gene, located on chromosome 15, is greater than 100 kb in size and the coding sequence contains 21 exons [135]. The product of this gene is a 180 kDa (1367 amino acids) precursor, which is proteolytically processed to yield α -chain (residues 1–707) and β -chain (residues 712–1337) [134, 136]. The mature IGF-1R is subsequently composed of two α - and two β -chains linked together via disulphide bonds to form $(\alpha\beta)_2$ homodimer (Figure 5.3.) which is heavily glycosylated [137-139]. Since glycosylation is critical for IR biosynthesis it is suggested the same is valid for IGF-1R, which has 16 potential N-linked glycosylation sites, 11 on the α -chain and 5 on the β -chain [134, 140].

α -Chain of IGF-1R consists of, from the N-terminus to the C-terminus, a leucine-rich repeat domain (L1 domain), a cysteine-rich region (CR domain) containing 22 cysteine

residues, a second leucine-rich domain (L2 domain), fibronectin type III domains (FnIII-1, and FnIII-2a) and an insert domain α (ID- α), which contains so-called α -CT peptide at the end [31, 141-144].

β -Chain is formed by, again from the N-terminus to the C-terminus, insert domain β (ID- β), fibronectin type III domains (FnIII-2b and FnIII-3), transmembrane sequence, tyrosine kinase domain and C-tail (Figure 5.3.) [31, 142-144].

Extracellular ectodomain of IGF-1R consists of two complete α -chains and 195 residues of each of β -chain (ID- β , FnIII-2b and FnIII-3 domains) and is responsible for binding of hormones. The transmembrane sequence (TM domain) is formed by the residues 906–929 of each β -chain, which continue by another 408 residues into the cytoplasm where form a juxtamembrane region (JM domain, residues 930–972), contains the phosphotyrosine binding sites for signaling molecule, tyrosine kinase domain (TK domain, residues 973-1229) followed by C-tail (residues 1230–1337) [3, 134, 139] (Figure 5.3.).

The 3D structure of the first three domains (L1-CR-L2) of the IGF-1R has been solved by X-ray crystallography [145]. In addition, several structures of the kinase domain have been also determined [146-149].

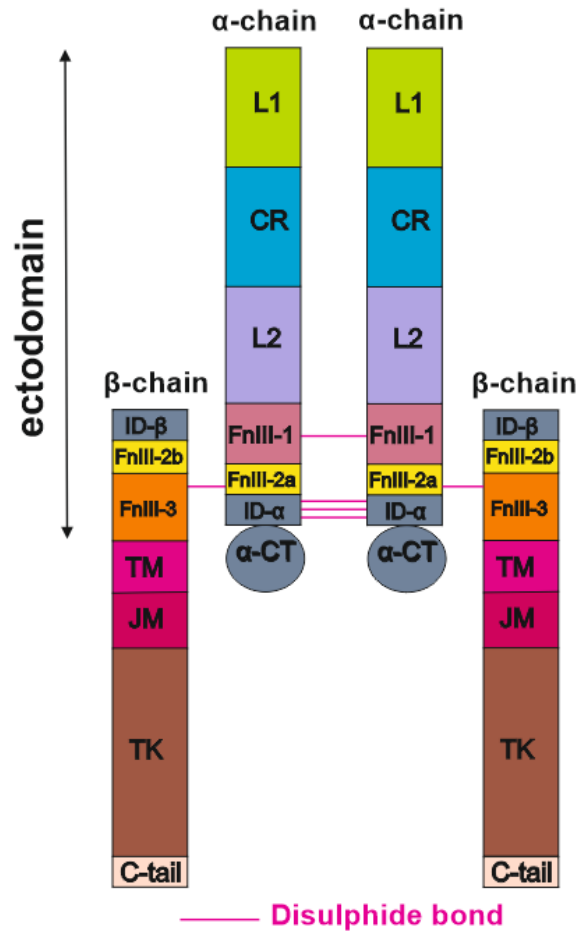


Figure 5.3. IGF-1R structure. The domain organisation of IGF-1R and the approximate location of the disulphide bonds are shown. Reproduced from Ref. [3].

5.4.2 Binding of hormones

Structural understanding of binding of IGF-1 and other hormones to the extracellular ectodomains of the IGF-1R is not completely understood. Until spring 2018 (see below), there was no complete crystal or any other structure of any of receptors of the tyrosine kinase superfamily [132]. However, the interaction of IGF-1 with its receptor is similar to the insulin/IR interaction where some structural advances were done.

It is widely acknowledged that insulin binding to the IR is mediated by two binding epitopes, so-called Site 1 and Site 2, which together form the complete binding pocket for the hormone. Site 1 (also known as ‘‘classical binding site’’) consists of L1 domain and α -CT peptide and binds insulin with a dissociation constant (K_d) of approximately 6.4 nM [150-154]. The first crystal structure of the complex of insulin sitting on L1 domain and α -CT peptide of IR representing Site 1 was published in 2013 [150] and further refined in 2014

[155]. This structure (Figure 5.4.) confirmed the results of previous studies suggesting the detachment of the C-terminus of the B-chain of insulin from the central core of the insulin molecule [156, 157].

The first crystal structure of IGF-1 in a complex with the receptor (hybrid construct consisting of L1-CR domains of IR and α -CT peptide of IGF-1R) showed that the binding of IGF-1 to the IR receptor binding Site 1 is very similar to the binding mode of insulin. However, this structure (Figure 5.5.) did not provide details about the arrangement of the C-domain of IGF-1 on the receptor [158].

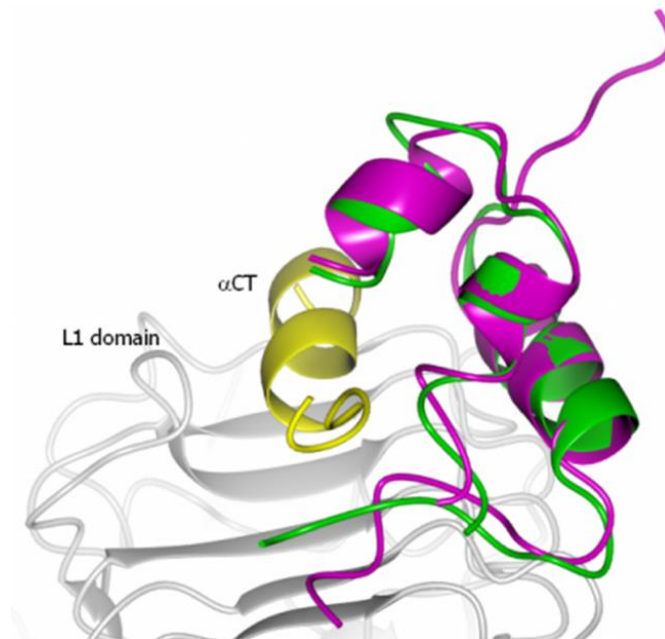


Figure 5.4. Crystal structure of the complex of insulin sitting on L1 domain and α -CT peptide of IR. An overlay of receptor bound insulin (in green) sitting on the L1 domain of the insulin receptor (in grey) and α -CT peptide (in yellow, 4OGA from Ref. [155]) with a crystal structure of AsnB26-insulin (in violet, 4UNG from Ref. [156]) mimicking the receptor-bound conformation of insulin. This complex represents so-called Site 1 of the interaction.

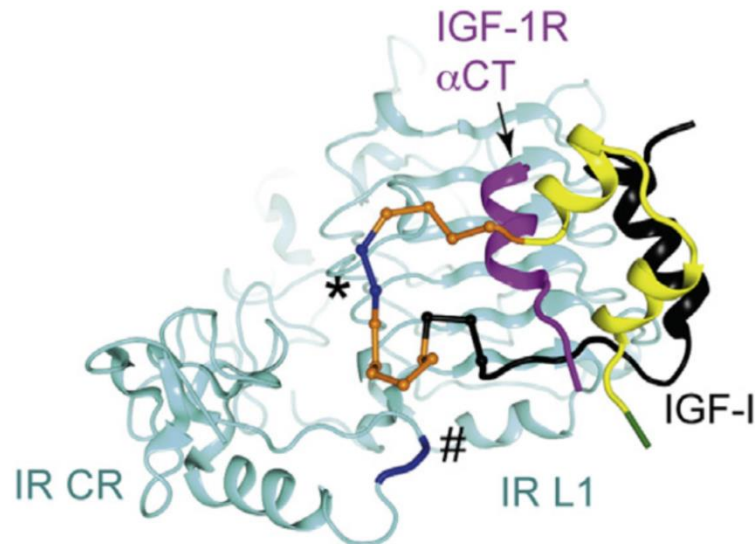


Figure 5.5. Crystal structure of the complex of human IGF-1 and receptor. IGF-1 bound to L1 and CR domains of IR (in light blue) and α -CT peptide of IGF-1R (in violet). A-domain of IGF-1 is in yellow, B-domain in black and E-domain in green. Orange and blue parts of the IGF-1 C-domains are shown in orange and were modelled in the complex because they were not visible in a crystal structure. Reproduced from the Ref. [158].

If the Site 1 of the IR receptor is relatively well defined from the available crystal structures (see above), the identity of Site 2 still remains unclear. It was proposed that it should be formed by FnIII-1 and FnIII-2 domains and its dissociation constant (K_d) value for insulin is about 400 nM [151, 153, 154, 159]. Therefore, only the simultaneous interaction of both Sites creates a high affinity complex with affinity about 0.2 nM [17].

The first structural insight into the second binding surface of insulin and IR was published early this year by Scapin et al. [160]. The Cryo-EM structure of the extracellular ectodomain of IR confirmed crystallographic data about Site 1-Site 1' interaction and proposed that Site 2 in insulin is formed only by residues CysA7, ThrA8 and B-chain residues B4-B10 (Figure 5.6.) and by Site 2' residues 495-498, 539-541 and 575 in IR FnIII-1 domain. This result is surprising especially because the extent of Site 2 in insulin is much more limited than predicted (see above). It is not excluded that some new discoveries will correct these findings.

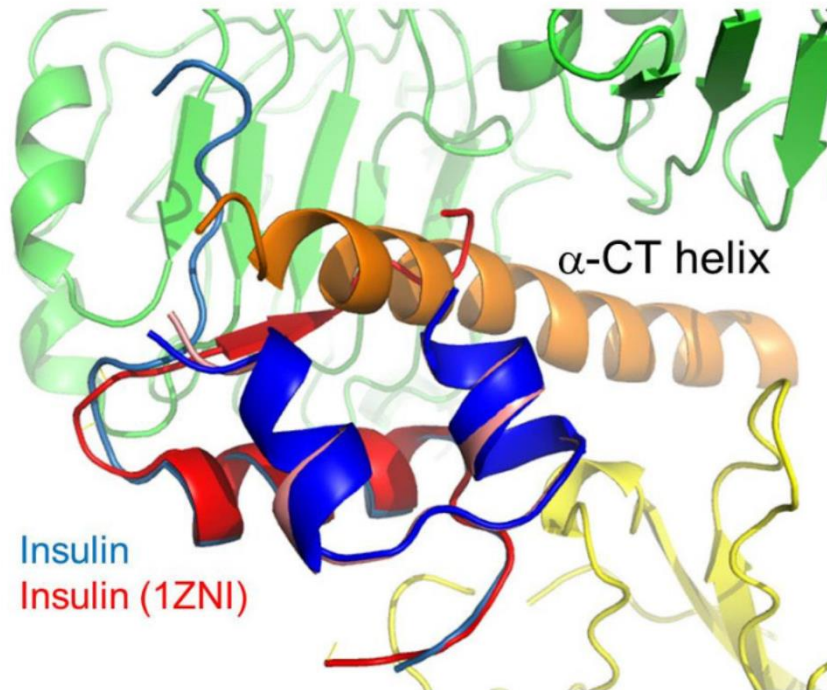


Figure 5.6. The CryoEM structure of the complex of insulin and insulin receptor. Bound insulin is in blue and is overlaid with free insulin in red (PDB1ZNI). Site 1' of the receptor is represented by L1 in green and α -CT peptide in orange. Site 2' of the receptor is in yellow and belongs to FnIII-1 domain. Reproduced from Ref. [160].

Interestingly, the first crystal structure of IGF-1 in a complex with IGF-1R was published almost simultaneously with Scapins's et al. Nature paper. Xu et al. [161] showed (Figure 5.7.) that IGF-1 binds to IGF-1R L1, CR and α -CT in an almost identical fashion as to IR/IGF-1R hybrid receptor (as shown above) [158]. They also identified contacts of IGF-1 with FnIII-2 IGF-1R domain, which were attributed to the supposed Site 2. However, the mutagenesis of FnIII-2 residues involved in this interaction did not confirm their importance for IGF-1 binding [161]. Hence, the authors conclude that FnIII-2 interaction with IGF1-R is likely non-physiological and probably caused by crystal soaking with the hormone.

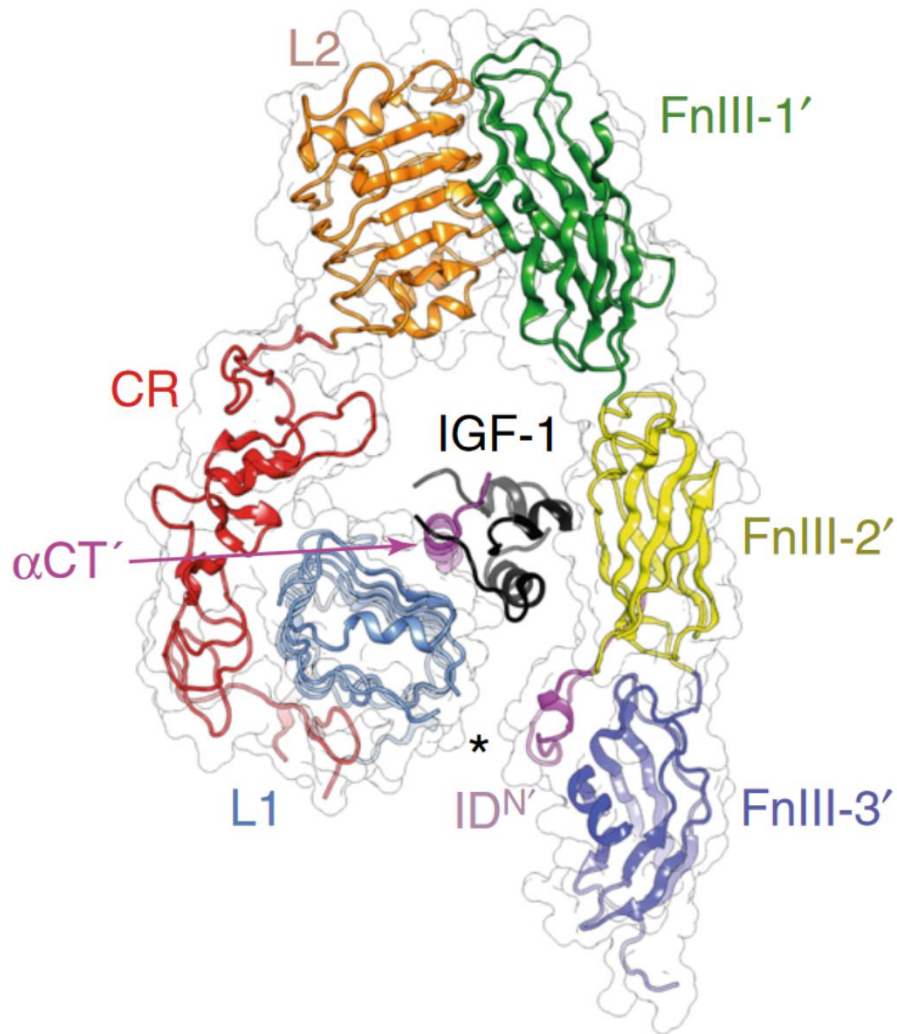


Figure 5.7. Crystal structure of a complex of IGF-1 and IGF-1R extracellular ectodomain. Reproduced from Xu et al. Ref. [161].

5.4.3 Signalling cascade

According to the classical model, hormone binding induces a conformational change of the receptor, leading to the activation of the TK domain via sequential autophosphorylation of three conserved tyrosine residues 1131, 1135, and 1136. In the unphosphorylated state, the receptor catalytic activity is at a low basal level [162-165].

Activation loop (A-loop), located at TK domain and containing the critical tyrosines 1131, 1135, and 1136, behaves as a pseudosubstrate that blocks the active site (closed configuration). After ligand binding, phosphorylation of Tyr 1135 and Tyr 1131 destabilizes the auto-inhibitory conformation of the A-loop and phosphorylation of Tyr 1136 and Tyr 1135 stabilizes the catalytically active conformation of TK domain [146].

The two most studied intracellular substrates of the IGF-1R tyrosine kinase are adaptor proteins Shc and the insulin receptor substrates (IRSs). They are well known to play important roles in subsequent signalling cascades [166].

Shc protein family contain four unique proteins (ShcA, B, C, and D) and their multiple splicing isoforms [167]. Their PTB (phosphotyrosine-binding) domain at the N-terminal region interacts with the phosphotyrosine P-Tyr⁹⁵⁰ (NPXY⁹⁵⁰ motif), located at the juxtamembrane region of the IGF-1R [168, 169]. Maximal phosphorylation of Shc is reached 5–10 min after IGF-1R stimulation [170].

The IRS protein family consists of four proteins (IRS1-IRS4) encoded by four genes (*Irs-1-4*), from which IRS-1 and IRS-2 are the best characterized. They bind to the IGF-1R in the same manner as Shc [169, 171] and their maximal phosphorylation is occurs 1–2 min after phosphorylation of the receptor [170]. In addition, IRSs can form high-molecular-mass complexes with a variety of proteins that modulate their availability to the IGF-1R [172, 173].

The two best-characterized pathways associated with IGF-1R activation are the MAPK (MEK) and PI3K pathways (see Figure 5.8.).

MAPK pathway is initiated by phosphorylation of Shc or IRSs adaptors through Grb2 molecule [174]. Grb2 is an adapter protein, which interacts with Shc or IRSs via its SH2 domain and with son of sevenless (Sos) protein, a guanine nucleotide exchange protein, via its two SH3 domains. Sos mediates the release of GDP and subsequent binding of GTP to the membrane-bound protein Ras [175, 176]. The activated Ras initiates a classical phosphorylation cascade, where Ras interacts with and translocates the serine/threonine protein kinase Raf to the plasma membrane, where Raf becomes activated. Activated Raf subsequently phosphorylates and thus activates one of the serine-threonine mitogen-activated protein kinases (MAPKs) known as MEK [177]. MEK in turn stimulates by phosphorylation the extracellular-related kinases (ERKs) ERK1 and ERK2, members of MAPKs group too, which are afterwards translocated to the nucleus [178]. ERK1/2 regulate cellular processes such as proliferation, differentiation and protein synthesis [179].

PI3K pathway begins by phosphorylation of the insulin receptor substrates (IRSs) and this interaction with a p85 regulatory subunit of PI 3-kinase (phosphoinositol 3-kinase) class I leads to activation of the catalytic subunit p110 of PI 3-kinase [180]. It has also been shown that PI 3-kinase can directly bind to the Y¹³¹⁶XXM motif of IGF-1R if this tyrosine is phosphorylated [181]. PI 3-kinase synthesizes the second messenger phosphatidyl inositol

(3, 4, 5)-triphosphate (PIP₃) from membrane-bound phosphatidylinositol (4,5)-bisphosphate (PIP₂). These phospholipids trigger phosphorylation of the Thr³⁰⁸ and Ser⁴⁷³ residues on Akt, a serine/threonine-specific protein kinase also known as protein kinase B, and activates this kinase [182, 183]. Activated Akt subsequently phosphorylates and thus inhibits several proapoptotic proteins [184-186], which leads to a suppression of apoptosis and causes survival effects. Other effects of Akt activation are metabolic and can result by the uptake of glucose to cells [166].

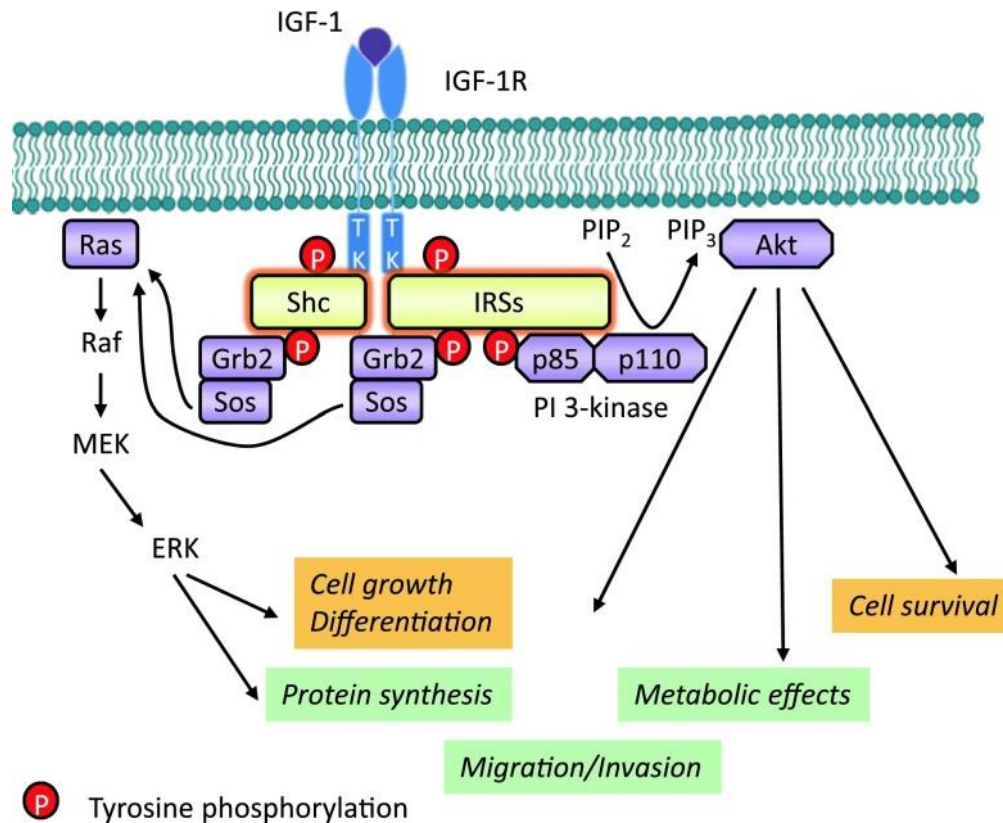


Figure 5.8. IGF-1R signalling pathways. Reproduced from Ref. [166].

5.4.4 IGF-1R internalization, recycling, and degradation

Ligand-activated receptors are normally downregulated by internalization allowing the cells to return to an unstimulated, basal state. This is also the case with IGF-1R. Only phosphorylated IGF-1R is internalized and, hence, the internalization is ligand-dependent process. IGF-1R is incorporated into specific clathrin- or calveolin-coated vesicles that transfer the receptor to the lysosome system for degradation [187].

After internalization, IGF-1R degradation is mediated by both the proteasome and lysosomal pathways [188]. Likewise, IGF-1R can be recycled to the plasma membrane [189]. However the details of these processes are still unknown [166].

Recent data demonstrated that IGF-1R is a substrate for ubiquitination and three E3 ligases, Mdm2 [190], Nedd 4 [191] and c-Cbl [187], have been shown to be involved in mediating ubiquitination of lysine residues of the IGF-1R. In Mdm2-mediated ubiquitination, β -arrestin functions as a molecular scaffold in bridging the ligase to the receptor [192]. Similarly Grb10 functions as an adapter protein in case of Nedd4-mediated IGF-1R ubiquitination [191].

5.4.5 *Nuclear IGF-1R*

Accumulating evidences indicate that intact, or proteolytically cleaved, fragments of IGF-1R (but also IR) translocate to the nucleus [193, 194]. It is suggested that IGF-1R nuclear import has a biological significance. Even though the mechanism is still unknown, it was shown to be mediated by SUMOylation [195, 196]. Interestingly, SUMOylation sites on lysine residues, located on the tyrosine kinase domain, are conserved among a variety of species and their mutagenesis arrested nuclear translocation and gene activation [196]. Further, one study showed that nuclear trafficking is mediated at least in part by clathrin and confirmed that both the α - and β -chains of IGF-1R translocate to the nucleus [194]. Moreover, other studies also demonstrate that ligand-mediated receptor phosphorylation is crucial for nuclear trafficking [197].

It was established that IGF-1R interacts, either directly or indirectly, with double-stranded DNA [195]. Notion that IGF-1R is actively translocated to the nucleus, where it can bind to enhancer regions of genomic DNA, prompted speculations about function of nuclear IGF-1R as a non-classical transcription factor [197]. A recent study has shown that nuclear IGF-1R binds and activates transcription factor LEF-1, a key regulator of the Wnt signaling cascade [198]. Several other receptor tyrosine kinase family members, including EGFR, ErbB, FGFR, MET, VEGF, were reported to travel to the nucleus and function as transcription factors as well [199].

5.4.6 *Hybrid receptors of IGF-1R*

Since the insulin receptors and the insulin-like growth factor 1 receptor have highly homologous structures, the one-half of IR and one-half of IGF-1R can heterodimerize, which

leads to the formation of IR/IGF-1R hybrid receptors. In addition, each of the IR isoforms is equally able to form hybrids with the IGF-1R [18, 200, 201]. The heterodimerization is believed to occur with a similar efficiency as homodimerization in close agreement with the random assembly model. Hybrid receptors are widely distributed in all mammalian tissues co-expressing both IR and IGF-1R receptors [202].

Several functional studies have demonstrated that both hybrid receptors, IR-A/IGF-1R and IR-B/IGF-1R, have basically the same binding properties. They bind IGF-1 with relatively high affinity, whereas insulin with much lower affinity. In other words, their binding characteristics are very similar to those of native IGF-1R [202-205]. Moreover, a recent study [206] clearly showed that both hybrid receptors are effectively activated only by IGF-1 and not by insulin, which can have important implications for development of insulin resistance or other health complications.

5.4.7 IGF-1R downregulation and diseases

Patients with mutations in IGF-1R, which would lead to a massive downregulation of the receptor, are rare, only a few patients with heterozygous mutations in the IGF-1R have been reported. The phenotypes of these patients appear to be variable. However, intrauterine and postnatal growth retardation, microcephaly and increased IGF-1 level, which can be low initially because of feeding problems, are consistent findings in these patients [79].

IGF-1R is more frequently connected with malignancy. Since IGF-1R plays a nontrivial role in the control of cell growth and works as antiapoptotic factor, his significant role in cancer diseases is not surprising. There are studies that have shown increased expression of IGF-1R or IGF-1 in cancer of breast, lung, thyroid, gastrointestinal tract, prostate and other tissues [207, 208]. Likewise, epidemiological studies connected higher serum IGF-1 levels with increased risk of pre-menopausal breast cancer and prostate, lung or colorectal cancers [80, 209-211] thus indicating that the hormone and its receptor could be promising anticancer therapeutic targets.

Treatment with agents that block IGF-1R function could enhance the efficacy of treatment by chemotherapeutic in case of resistance [2]. Several experimental strategies have been tried to block IGF-1R function, or expression. First of all, reduction in the number of IGF-1R receptors was used, by antisense oligonucleotides and short-interfering RNA (siRNA) [2, 212], or by directly increasing receptor degradation [213, 214]. The majority of these systems were unsuitable for transfer to the clinical trials because of problems with

application to human treatment or with specificity, so new studies were focused on clinically more suitable methods for blocking the IGF-1R function.

An option is blocking of IGF-1R function by inhibiting IGF-1 binding using monoclonal antibodies, which have a high affinity for the hormone binding domain of the receptor. Interestingly, in some cases the antibody also reduces a population of IGF-1R by enhancing the rate of receptor internalization. The antibodies were tested especially on model systems of human tumour cells [2]. The initial studies used mouse monoclonal antibodies [215-217], and later, humanized [218] or fully human [219] antibodies have been also prepared and tested. Some of these antibodies had significant activity and thus were tested in clinical trials. Despite the promising results of the preclinical studies, the clinical results have been disappointing. Responses were observed only in cases of relatively rare tumours, Ewing's sarcoma [220-222], non-small cell lung carcinoma [223], adrenal cortical sarcoma and carcinoma [224] and even in these cases not all patients benefited from IGF-1R inhibition [166, 225].

Another option is the inhibition of IGF-1R function by using tyrosine kinase inhibitors that inhibit the enzymatic activity of the receptor. The development of specific inhibitors of the IGF-1R tyrosine kinase is a challenge owing to structural similarities with the insulin receptors. Tyrosine kinase domain is the most homologous part (84 %) of the IGF-1R and insulin receptors [134]. First large group of these compounds have been designed to compete for the ATP binding site on the tyrosine kinase. However, some of these inhibitors bind to the insulin receptor too. In addition, they also prevent IGF-1R downregulation [166].

A distinct category of IGF-1R kinase inhibitors is the cyclooligan picropodophyllin (PPP) and its analogues. PPP blocks phosphorylation of Tyr1136 residue, within the kinase domain A-loop, while sparing the two others (Tyr 1131 and Tyr 1135) [226]. Its great advantage, in contrast to ATP binding site inhibitors, is that PPP also triggers IGF-1R downregulation [227].

Inhibitors binding to ATP site of IGF-1R kinase [228-234] and PPP analogues [235-242] were used in a model of immunocompromised mice and human tumour xenografts. Several of these inhibitors have been tested in clinical trials too [243, 244].

Finally, overexpression of inhibitory forms of IGFBPs or antibodies against both IGF-1 and IGF-2, which mimic natural IGFBPs action, have also been suggested as potential therapeutic strategies. A great advantage of these strategies should be the lack of interference

with insulin signalling [245, 246]. Also some peptide analogues of IGF-1 have been shown to be able to inhibit the growth of cell lines [212].

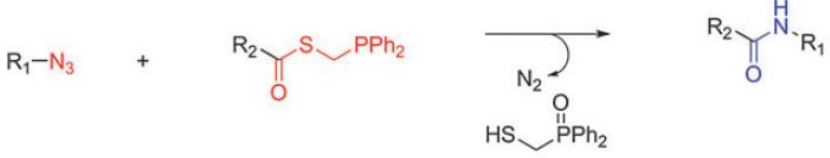




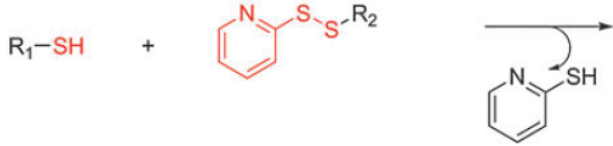

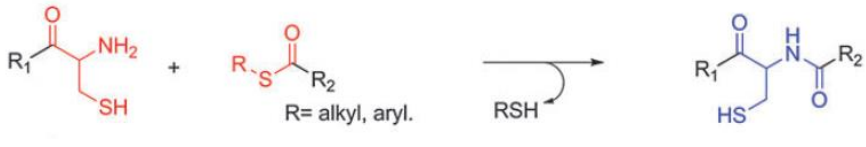


However, the results of clinical trials were not satisfactory, because of either the toxicity of the TK-targeting drugs or because of an increasing overlap and takeover of IGF-1R signalling pathways by the IR [247]. The lack of progress in addressing one of the key hallmarks of cancer underlines the need for new anticancer therapies that would exploit alternative, and specific, targets of the insulin/IGF system. Here, a high-affinity/no-efficacy IGF-based IGF-1 analogue, i.e., selective antagonist of the IGF-1R, should represent a promising new strategy for combating IGF-1R-related malignancies.

5.5 Click chemistry

Term “click reaction” was first used by K. B. Sharpless in 2001 for conjugation reactions, which are stereospecific, high yielding, wide in scope and simple to perform under biologically benign conditions (aqueous solution, ambient temperature, and near physiologic pH) [248]. Several reaction types have been identified that meet these criteria and thus these organic reactions are used as ligation tools for the synthesis of bio-conjugates, such as strain-promoted azide–alkyne cycloaddition (SPAAC), thiol–ene reaction, Diels–Alder reaction, oxime ligation, native chemical ligation and many others (see table 5.4.). However the most powerful and favourite one to date is the Cu^I-catalysed variant of the Huisgen 1,3-dipolar cycloaddition of azides and alkynes (CuAAC) [249].

Table 5.4. Summary of click reactions that are used to synthesize bioconjugates [249].

Reaction	Reaction scheme
CuAAC	
SPAAC	
Non-traceless Staudinger Ligation	

Traceless Staudinger Ligation	R_1-N_3 + $R_2-C(=O)S-CH_2-PPh_2$ 
Tandem crD-A	R_1-N_3 + $R_2-O-C(=O)-C_5H_3(F_3C)-O$ 
Thiol-ene	R_1-SH + $CH_2=CH-R_2$ 
Thiol-yne	R_1-SH + $C\equiv C-R_2$  R_1-SH + $CH_2=CH-CH(R_2)-EWG$ 
Thiol-Pyridyl Disulphide	R_1-SH + $Pyridine-S-S-R_2$ 
Thiol-halogen Ligation	R_1-SH + $X-CH_2-C(=O)-R_2$ X = Br, I. 
Native Chemical Ligation	$R_1-C(=O)-CH_2-NH_2$ + $R-S-C(=O)-R_2$ R = alkyl, aryl. 
Thiazolidine Ligation	$R_1-C(=O)-CH_2-NH_2$ + $H-C(=O)-R_2$ 
Oxime Ligation	$R_1-C(=O)-R$ + H_2N-O-R_2 R = H, alkyl. 

Diels-Alder Reaction	<p>diene: R_1-furan, R_1-cyclopentadiene. dienophile: R_2-maleic anhydride.</p>
Hetero Diels-Alder Reaction	<p>X = -P(=O)(OEt)2, pyridine.</p>
Inverse Electron Demand Diels-Alder Reaction	

5.5.1 Cu^I catalysed 1,3-dipolar cycloaddition of azides and alkynes

The Cu^I catalysed 1,3-dipolar cycloaddition of azides and alkynes, also called the copper-catalyzed azide-alkyne cycloaddition (CuAAC), is reaction of azides and alkynes to form 1,2,3-triazoles (Figure 5.9.) was first published by Meldal et al [250] and further developed by the same lab and by Sharpless' lab [251, 252].

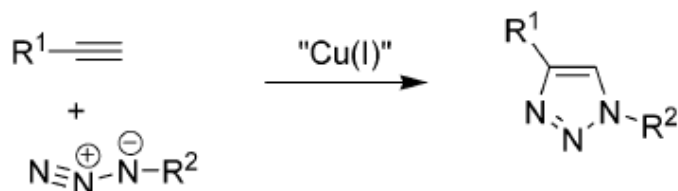


Figure 5.9. Reactants and product of the Cu^I catalysed 1,3-dipolar cycloaddition [253].

These 1,2,3-triazole heterocycles have characters of stable and rigid linking units that can mimic some properties of the peptide bond, but are resistant to hydrolytic cleavage (see

Figure 5.10.). Few structural differences between triazoles and amide bonds of course exist. First of all, the extension of one extra atom in the triazole backbone leads to a 1.1 Å longer distance R¹–R² in comparison to the typical amide bond. Moreover, triazoles also possess a much stronger dipole moment than an amide bond [254].



Figure 5.10. Structure of electronic and topological similarities of amides and 1,2,3-triazoles [254].

A long list of applications suggests that Cu^I-catalysed alkyne–azide coupling affords triazoles mostly in good yields under varying conditions and thus it is accepted as a robust reaction [255]. Although CuAAC is described as almost quantitative conversion in organic synthesis, however, in the case of peptide conjugations (as well as polysaccharide conjugation), the secondary structures of peptide chains may hinder the efficiency of the reaction. Increased temperature and microwave irradiation have been demonstrated to increase the reaction rates and the yields [249, 256, 257].

Another common problem of this reaction is the oxidation of Cu^I to Cu^{II}, which can be caused by air oxygen. Degassing and protection by an inert gas, as well as the addition of reducing agent like sodium ascorbate are routine strategies to avoid the Cu^I oxidation. In addition, the use of rate-accelerating ligands is a common procedure to limit copper oxidation. These ligands are capable of binding to Cu^I by forming a stable chelate and thus prevent it from oxidation [258].

6 Research Aims

Project 1. Fully synthetic IGF-1 analogues

- Development of a new methodology for the total chemical synthesis of IGF-1 analogues.
- Preparation of a series of fully synthetic IGF-1 analogues with modifications at the positions Arg36 and Arg37 in the IGF-1 C domain and at the position Met59 of the A domain.
- Biochemical characterization of new analogues for binding to cognate receptors (IR-A and IGF-1R) and for inducing autophosphorylation of these receptors.

Project 2. Recombinant IGF-1 analogues

- Development of a new methodology for a recombinant production of IGF-1 and its analogues in *E. coli*.
- Preparation of recombinant IGF-1 analogues with modifications at the positions Asp45, Glu46 and His49 in the IGF-1 A domain.
- Biochemical characterization of new analogues for binding to cognate receptors (IR-A and IGF-1R) and for inducing autophosphorylation of these receptors.

7 Publications

Publications included in the dissertation thesis

1. **Macháčková K**, Collinsová M, Chrudinová M, Selicharová I, Pícha J, Buděšínský M, Vaněk V, Žáková L, Brzozowski AM and Jiráček J. *Insulin-like Growth Factor I Analogs Clicked in the C Domain: Chemical Synthesis and Biological Activities*. J. Med. Chem., 2017, **60**(24), pp 10105-10117.
2. Pícha J, Buděšínský M, **Macháčková K**, Collinsová M and Jiráček J. *Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides*. J. Pept. Sci., 2017, **23**(3), pp. 202-214.
3. Hexnerová R, Křížková K, Fábry M, Siegllová I, **Kedrová K**, Collinsová M, Ullrichová P, Srb P, Williams C, Crump MP, Tošner Z, Jiráček J, Veverka V and Žáková L. *Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain*. J. Biol. Chem., 2016, **291**(40), pp 21234-21245.
4. **Macháčková K**, Chrudinová M, Radosavljević J, Potalitsyn P, Křížková K, Fábry M, Selicharová I, Collinsová M, Brzozowski AM, Žáková L, and Jiráček J. *Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation*. Biochemistry, 2018, **57**(16), pp 2373–2382.

7.1 Insulin-like Growth Factor 1 Analogues Clicked in the C Domain: Chemical Synthesis and Biological Activities

Background:

The central biological importance of IGF-1 and its involvement in many relevant human pathologies led to the use of IGF-1 analogues as molecular tools for the studies of interactions with IGF-1 receptor (IGF-1R) and as potential new therapeutics. The recombinant production of insulin and IGFs in *E. coli*, or in yeast, is an obvious choice [259, 260], but it does not allow a facile incorporation of non-coded amino acids and unusual structural motifs. Such unusual structural motifs can offer new possibilities to structure–activity studies of the hormones and their rational modifications. Here, the organic synthesis offers more possibilities and variability than recombinant production.

The biological aspect of this work was focused on the role of a pair of arginines at the positions 36 and 37 located within IGF-1 C domain. This motif was previously identified by the others as important for binding to the IGF-1 receptor (IGF-1R) [20] and to IGF binding protein 1 (IGFBP-1) [21] and their mutations to Glu36 and Glu37 resulted in a partial antagonism towards IGF-1R [261]. Moreover, it was suggested [13] that the peptide bond between Arg36 and Arg37 can be cleaved in vivo being a mechanism of hormone's deactivation but hypothetically also a maturation of the prohormone to its active form.

Summary of the results:

Using the solid-phase peptide synthesis and biomimetic formation of a native pattern of disulphides we prepared a two-chain IGF-1 analogue (disconnected between Arg36 and Arg37) and three single-chain IGF-1 analogues with a connecting 1,2,3-triazole moiety between, or in, sites 36 and 37 to probe the feasibility of the chemical synthesis of IGF-1 analogues and to test the impact of these modifications on the biological activities of these analogues. We found that the solid phase synthesis of two shorter IGF-1 fragments and their subsequent connection by the triazole moiety is a convenient method to overcome difficulties with the synthesis of too long peptide chains. For this, we also prepared a series of N-Fmoc-protected azido amino acids and we optimized their incorporation to peptide chains. Biomimetic formation of a native pattern of disulphides enabled to avoid laborious preparation of peptide chains with orthogonally protected cysteine. These new synthetic

IGF-1 analogues are unique examples of disulphide bonds rich proteins with intra main-chain triazole links.

The binding affinity of two-chain IGF-1 was severely reduced indicating that the cleavage of the 36-37 peptide bond is more likely a degradation than an activation step. The intra-chain triazole moiety was relatively well tolerated in IGF-1R binding and its activation and the resulting analogues did not show any signs of antagonism. The methodology reported here presents a convenient synthetic platform for the design and production of new analogues of this important human hormone with non-standard protein modifications.

My contribution:

I developed a new methodology for the total chemical synthesis of IGF-1 analogues and prepared all analogues with the triazole link in the main chain. I participated in the measurement of their binding affinities and in the preparation of the manuscript.



Insulin-like Growth Factor 1 Analogs Clicked in the C Domain: Chemical Synthesis and Biological Activities

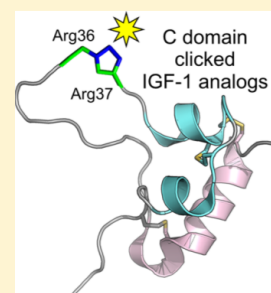
Kateřina Macháčková,[†] Michaela Collinsová,[†] Martina Chrudinová,[†] Irena Selicharová,[†] Jan Pícha,[†] Miloř Buděřinský,[†] Václav Vaněk,[†] Lenka Žáková,[†] Andrzej M. Brzozowski,[‡] and Jiř Jiráček^{*,†}

[†]Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

[‡]York Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York YO10 5DD, United Kingdom

Supporting Information

ABSTRACT: Human insulin-like growth factor 1 (IGF-1) is a 70 amino acid protein hormone, with key impact on growth, development, and lifespan. The physiological and clinical importance of IGF-1 prompted challenging chemical and biological trials toward the development of its analogs as molecular tools for the IGF-1 receptor (IGF1-R) studies and as new therapeutics. Here, we report a new method for the total chemical synthesis of IGF-1 analogs, which entails the solid-phase synthesis of two IGF-1 precursor chains that is followed by the Cu^I-catalyzed azide–alkyne cycloaddition ligation and by biomimetic formation of a native pattern of disulfides. The connection of the two IGF-1 precursor chains by the triazole-containing moieties, and variation of its neighboring sequences (Arg36 and Arg37), was tolerated in IGF-1R binding and its activation. These new synthetic IGF-1 analogs are unique examples of disulfide bonds' rich proteins with intra main-chain triazole links. The methodology reported here also presents a convenient synthetic platform for the design and production of new analogs of this important human hormone with non-standard protein modifications.



INTRODUCTION

Human insulin and related insulin-like growth factors 1 and 2 (IGF-1 and IGF-2) are small protein hormones, with a key impact on basal metabolism, growth, development,^{1–3} lifespan,^{4,5} as well as being implicated in diabetes and cancer.^{6,7} Hence their central biological importance and involvement in many relevant human pathologies require efforts toward the development of specific analogs of these hormones for studies of the interactions and signaling with respective insulin- and IGF-receptors (IR, IGF-1R) and their applications as new therapeutics.^{3–9} The recombinant production of insulin and IGFs in *E. coli*, or in yeast, was an obvious choice,^{10–12} but did not easily allow the incorporation of non-coded amino acids and unusual structural motifs, which could facilitate more directed structure–activity studies of the hormones and their rational modifications. Moreover, the total chemical synthesis of these proteins, especially of IGF-1 and -2, remains a very challenging task.¹³

The single-chain organization, fold, and disulfide bond pattern of IGF-1/IGF-2 is similar to proinsulin, a prohormone of insulin, but they are significantly longer (70, 67, and 86 residues, respectively) than both insulin chains (21 and 30 amino acids). Therefore, their solid-phase synthesis is challenging and inefficient as demonstrated by only few early reports on the total chemical synthesis of IGF-1 analogs.^{14–16} Nevertheless, elegant studies employed a combination of shorter protein fragments by a native ligation for the synthesis

of human proinsulin,¹⁷ single chain “ester” insulin,^{18,19} and IGF-1 analog.²⁰ The complexity of the synthesis of these hormones is highlighted further by only one report on the total chemical synthesis of IGF-2.²¹

Therefore, our aim here was to develop a new methodology for the chemical synthesis of IGF-1 analogs as a feasible alternative for the production of single-chain insulin-like hormones, and one that would allow easy incorporation of non-standard amino acids or other unusual structural motifs.

As the solid-phase synthesis of a single, 70-amino acid chain of IGF-1 is inefficient, in this work we combined the solid-phase synthesis of two shorter precursor IGF-1 fragments with their ligation by a Cu^I-catalyzed 1,3-dipolar cycloaddition of azides and alkynes,^{22–24} followed by the biomimetic formation of the native pattern of disulfide bridges. Cu^I-catalyzed 1,3-dipolar cycloaddition of azides and alkynes is a biorthogonal reaction, also called a “click” chemistry, which leads to the formation of 1,4-disubstituted 1,2,3-triazoles. This reaction has become widely used²⁵ in organic, medicinal, and, especially, peptide chemistry,²⁶ as 1,2,3-triazole is a motif with the structural and electronic characteristics similar to those of the peptide bond.^{27,28}

The synthesis of insulin and insulin-like proteins is difficult due to the necessity to form three disulfide bridges. Therefore,

Received: September 6, 2017

Published: November 27, 2017

a number of sophisticated methods of insulin synthesis have been developed, usually with an orthogonal protection of cysteines (e.g., in refs 29 and 30 and reviewed in ref 31) and a directed, stepwise formation of disulfides. A similar strategy was developed for other proteins from the insulin family, such as relaxins, DILP, INSL peptides,^{32–36} and cone snail insulin.³⁷ An alternative, and straightforward, method for the formation of disulfides starts with the simultaneous deprotection of all six cysteines (from e.g. *S*-sulfonate protected forms) and pairing by a slow air oxidation. This approach takes advantage of the intrinsic biomimetic ability of insulin chains to adopt a natural conformation and a correct coupling of disulfides.^{18,38–43} In this study, we used the strategy of biomimetic formation of disulfides for the synthesis of IGF-1 analogs.

The biological functionality aspect of this work was focused on the role and modifications of the IGF-1 C domain, specifically, at its Arg36 and Arg37 sites. The C domain of IGF-1 (residues 30–41, Figure 1) was found to be important for its

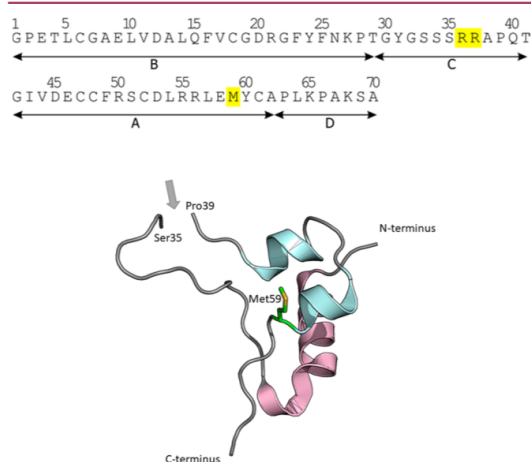


Figure 1. The primary sequence and the 3D fold of human IGF-1. The primary sequence of human IGF-1 (one-letter code, N-terminus is in the left, C-terminus is in the right, upper panel); residues mutated in this study (Arg36, Arg37 and Met59) are highlighted in yellow; disulfide bridges: 47–52, 48–6, and 61–18. The organization of IGF-1 into B, C, A, and D domains is shown below the sequence by arrows, and the 3D fold of its main chain is in the lower panel (1GZR.PDB).⁵¹ The gap in the protein chain (marked by a gray arrow) reflects a typical lack of connectivity between Arg36–Ala38; Met59, mutated in one of the analogs is shown in green, IGF-1 residues 39–66 are in light blue, residues 3–35 are in pink; residues 67–70 and 1–2 were disordered.

functionality,^{44–46} as the double Arg36–Arg37 → Ala36–Ala37 mutation severely impaired binding of this mutant to the IGF-1 receptor (IGF-1R)⁴⁷ and to IGF binding protein 1 (IGFBP-1), which contributes to the regulation of IGF-1 bioavailability.⁴⁸ In addition, a double Arg36–Arg37 → Glu36–Glu37 substitution also indicated a partial antagonism of this analog, that is, its still effective binding to IGF-1R, but an impaired receptor signaling⁴⁹ and tumorigenesis.⁵⁰

Interestingly, the connectivity, that is, well-defined electron density for residues Arg36–Arg37–Ala38, and, generally, for this central part of the C domain, was frequently missing in several crystal structures of human free, and IGFBP-bound, IGF-1

(Figure 1).^{51–53} It was postulated then that the pair of these basic arginine side chains may constitute an *in vivo* proteolytic cleavage site. Although the presence of Arg36 and Arg37 in the IGF-1 C domain is important for an efficient binding to IGF-1R,⁴⁷ it was not known whether the cleavage of the Arg36–Arg37 peptide bond impairs hormone binding to IGF-1R or if the resulting two-chain IGF-1 is still able to effectively activate IGF-1R. For this reason, we produced a two-chain IGF-1 analog here as well, which is disconnected between sites Arg36 and Arg37.

Together, we designed, synthesized, and characterized one two-chain IGF-1 analog and three single-chain IGF-1 analogs with a connecting triazole bridge between, or in, sites 36 and 37 to probe the feasibility of their synthesis and to test the impact of these modifications on the biological activities of these analogs. In one of the analogs, Met59 was mutated to Nle59 to prevent oxidation of Met. The *in vitro* functional properties of the analogs were characterized by the determination of their binding affinities to the IGF-1R, insulin receptor isoform A (IR-A), and their receptor-activation potencies. The IR-A receptor was used here as an IR isoform that, in contrast to its IR-B isoform, is a relatively good binder of IGF-1 and, especially, of IGF-2.^{12,54,55}

RESULTS

Synthesis. The primary structures of peptide precursors and analogs are shown in Figure 2, and the respective synthetic pathways are overviewed in the Supporting Information Figures S1–S4.

Two-Chain IGF-1 Analog. First, a two-chain human IGF-1 was prepared, referred to as analog 1, which has a native IGF-1 amino acid sequence, with the same pattern of the disulfide bonds, but it has a split Arg36–Arg37 peptide bond. Hence, analog 1 consists of chain 1–36 and chain 37–70, with all N-termini and C-termini being in the free amine or carboxylic acid forms, respectively (Figure 2).

The individual chains of the analog were prepared by a standard solid-phase synthesis, and the free SH groups of cysteines were converted to *S*-sulfonates (chains 1a and 1b, Figure 2) immediately after the cleavage of Trt protecting groups (Figure S1). The chains were recombined (i.e., disulfides formed) to give analog 1 with approximately 0.6% yield (0.2 mg) by a procedure described previously.^{42,43}

Single-Chain IGF-1 Analogs. We designed the strategy of the solid-phase peptide synthesis of two shorter IGF-1 chains with β -azido-alanine or propargyl glycine at their C- or N-termini, respectively, followed by their ligation by the Cu^I-catalyzed azide–alkyne cycloaddition.^{22–24} The final synthetic step consisted of the simultaneous biomimetic formation of the native pattern of all three disulfide bridges, as in analog 1.

The analog 2 was synthesized first (Figures 2 and S2). The synthesis started with the preparation of the chain 2a, in which the C-terminal Arg36 was replaced by β -azido-alanine carboxamide, and the chain 2b, in which N-terminal Arg37 was replaced by propargyl glycine. Both chains were “clicked”, and the single-chain (Acm-protected) precursor 2c was purified by RP-HPLC. Finally, the Cys-protecting acetamidomethyl (Acm) groups were cleaved, and the analog 2 was allowed to fold by air-induced oxidation.

Next, we synthesized the similar analog 3 from the chains 2a, 3a, and 3b (Figures 2 and S3). The only difference between analogs 2 and 3 was that Met59 was substituted in 3 for norleucine (Nle) to prevent the observed oxidation of Met59

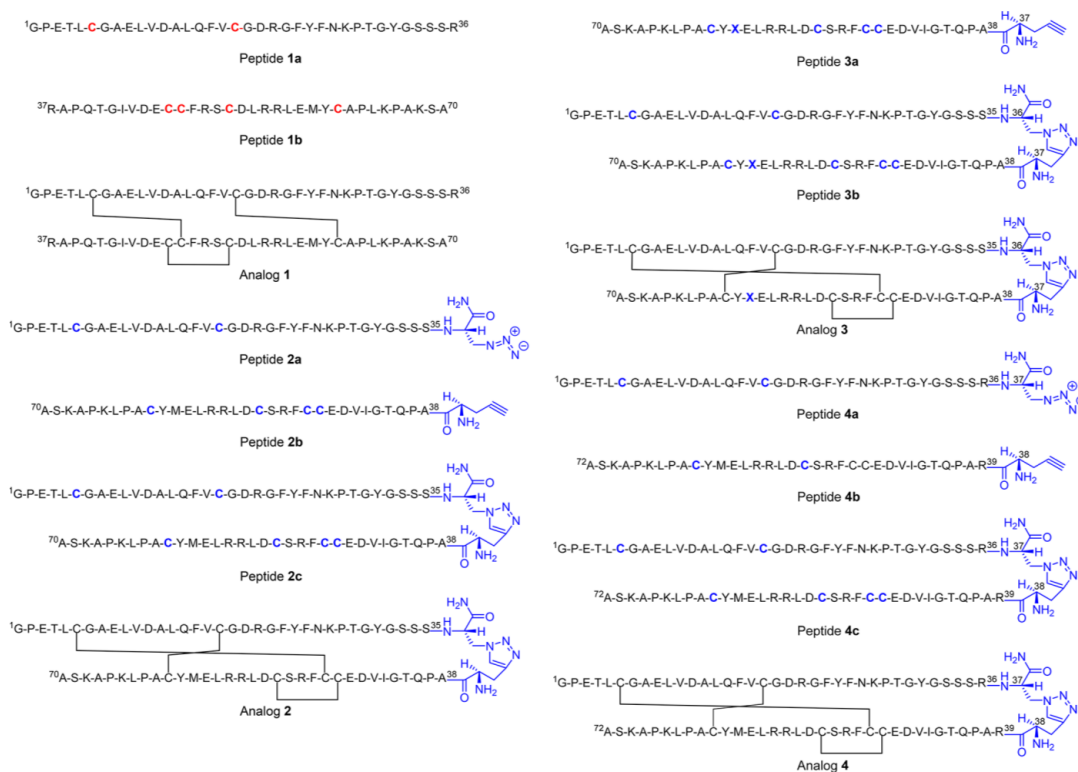


Figure 2. Primary structures of peptide precursors and analogs synthesized in this work. Amino acids are shown in one-letter code. The connecting lines between Cys residues represent disulfide bridges. Unnatural structural features in peptides are highlighted in red or in blue. Red C in peptides **1a** and **1b** represents Cys residues protected by S-sulfonate groups. Blue C in peptides **2a**, **2b**, **2c**, **3a**, **3b**, **4a**, **4b**, and **4c** represents Cys residues protected by AcM groups. Blue X in peptides **3a**, **3b**, and analog **3** is norleucine. Azido or propargyl amino acids in **2a**, **2b**, **3a**, **4a**, and **4b** and connecting triazole-containing moieties and **2c**, **2b**, **3b**, **3c**, **4c**, and **4** are shown in blue as well. ¹G and ³⁷R in peptides **1a**, **1b**, and analog **1** are N-terminal residues, and ³⁶R and ⁷⁰A in the same peptides are C-terminal residues. ¹G is N-terminal glycine in peptides **2a**, **2c**, **3b**, **4a**, **4c**, and analogs **2–4**. ⁷⁰A is C-terminal alanine in peptides **2b**, **2c**, **3b**, **3c**, and analogs **2** and **3**. ⁷²A is C-terminal alanine in peptides **4b**, **4c**, and analog **4**. All terminals of peptides and analogs are in the free forms.

side chain. Finally, we designed and made analog **4** (from the chains **4a**, **4b**, and **4c**, Figures 2 and S4), in which Arg³⁶ was preserved, the triazole link was formed between sites 37–38, with an additional Arg at the position 39; these modifications resulted in an analog that is two amino acids longer than native IGF-1.

In general, the chains with propargyl glycine at the N-terminus (**2b**, **3a**, and **4b**) gave better yields (15–18%) than the azido chains **2a** and **4a** (6–13%). The preparation of the precursor **2a** was hampered by the formation of some byproducts, which were identified as aspartimides (racemized) at positions Asp²⁰ and Arg²¹ (see analyses and model experiments described in Supporting Information). This problem was largely solved by the use of Fmoc-Asp(OEpe) instead of Fmoc-Asp(OtBu), and the yield improved from 6% to 10%, respectively. We also observed the generation of a byproduct during the synthesis of the peptide **2b**, due to the oxidation of Met⁵⁹ to methionine sulfoxide. This prompted synthesis of analog **3**, for which the chain **3a** with Met⁵⁹Nle mutation was used.²⁷ Although this resulted in a better purity of the crude product **3a**, the yields of HPLC purified peptides **2b**

and **3a** were similar (18% and 15%, respectively). Therefore, the native Met⁵⁹ was retained in peptide **4b**, with only a marginal oxidation of the final product.

Our first attempts (Figure S1) at the click ligation of the azido and alkyne precursor chains were performed as described previously,^{42,43} with cysteines protected as S-sulfonates. However, we have not succeeded to isolate the desired products here: None of the reaction conditions used (CuBr/TBTA/DIPEA, CuBr/THPTA/DIPEA, CuI/DIPEA, CuSO₄/ascorbic acid/THPTA, or CuSO₄/ascorbic acid in different organic or aqueous solvents with or without chaotropic agents) was successful, that is, compatible with the use of S-sulfonates.

Therefore, the strategy for an efficient disulfide pairing was changed, and all subsequent precursor peptides were synthesized with cysteines protected with AcM groups. Different conditions for the click reactions were investigated, but the only effective setup consisted of an excess of CuSO₄/ascorbic acid in tBuOH/water, and this methodology^{56,57} was employed for the preparation of peptide chains **2c**, **3b**, and **4c**. The effect of different molar ratios of chain precursors for the synthesis of **2c** and **4c** was also tested, and it was observed that

Table 1. The Receptor-Binding Affinities of Analogs 1–4 Reported in This Work^{a,d}

Analog	K_d	SD [nM] (n) for human IGF-1R in mouse fibroblasts	Relative binding affinity for human IGF-1R [%]		K_d	SD [nM] (n) for human IR-A in IM-9 lymphocytes	Relative binding affinity for human IR-A [%]	
Human IGF-1	0.24	0.10 (5) ^b	100	42	23.8	11.5 (3) ^{c,d}	1.0	0.5
	0.24	0.12 (5)	100	50				
	0.12	0.01 (5)	100	8.3				
Human insulin	291.8	54.3 (3) ^{b,e}	0.08	0.01	0.24	0.05 (5) ^c	100	21
					0.20	0.04 (5)	100	20
					0.12	0.04 (3)	100	33
					0.32	0.09 (4)	100	28
1	6.74	2.62 (3) ^e	3.56	1.38**	48.2	5.0 (3) ^f	0.41	0.04*
2	1.15	0.45 (4) ^g	20.8	8.1**	48.3	7.5 (3) ^h	0.25	0.04*
3	5.04	2.31 (3) ^d	4.76	2.18**	113	53 (3) ^d	0.21	0.09*
4	0.49	0.12 (3) ⁱ	24.5	6.0***	75.1	6.2 (3) ^j	0.43	0.04*

^aThe values of K_d and relative binding affinities (relative receptor-binding affinity is defined as (K_d of human insulin or IGF-1/ K_d of analog) \times 100) of human IGF-1, insulin, and the analogs were determined for human IGF-1R in membranes of mouse fibroblasts and human IR-A in membranes of human IM-9 lymphocytes. Asterisks indicate that binding of the ligand differs significantly from that of IGF-1 (* p < 0.05; ** p < 0.01; *** p < 0.001). ^bFrom Vikova et al. (ref 57). ^cFrom Krizkova et al. (ref 55). ^dRelative to human insulin K_d value of 0.24 0.05 (n = 5). ^eRelative to human IGF-1 K_d value of 0.24 0.10 (n = 5). ^fRelative to human insulin K_d value of 0.20 0.04 (n = 5). ^gRelative to human IGF-1 K_d value of 0.24 0.12 (n = 5). ^hRelative to human insulin K_d value of 0.12 0.04 (n = 3). ⁱRelative to human IGF-1 K_d value of 0.12 0.01 (n = 5). ^jRelative to human insulin K_d value of 0.32 0.09 (n = 4).

the cycloadditions at equimolar ratio of the precursor chains gave lower yields than the reactions with 2-fold molar excess of the azido chains. However, the optimum ratio of azido and propargyl precursors for **3b** was 1:1. Moreover, the use of more than 0.5 μ mol (about 2 mg) of the yield-limiting propargyl-containing chain was ineffective. As a result, several click reactions for each analog were performed to accumulate a sufficient amount of the material for their functional testing. The yields of optimized click reactions were typically about 16%, 18%, and 7% for **2c**, **3b**, and **4c**, respectively, after HPLC purification.

The final part of the synthetic work was focused on the optimization of an efficient, and proper/native, formation of the disulfide bonds in single-chain precursors **2c**, **3b**, and **4c**. Initially, an access of iodine in an acetic/hydrochloric acid mixture was used to cleave the AcM groups from Cys side chains in **2c** (Figure S2).³⁰ However, this was ineffective, as the majority of the AcM groups remained uncleaved, even after 4 h. Therefore, silver triflate was employed for the removal of the AcM groups, without the purification of an intermediate product. Subsequently, the free -SH groups were "liberated" from the Ag⁺ salt by the gaseous H₂S, and the precipitated Ag₂S was removed by centrifugation. We followed here our successful application of such an approach for the removal of copper from synthesized compounds after click reactions.^{58–60} Finally, the dilution of the supernatant, and the removal of an excess of the reductive H₂S gas, was followed by a slow air oxidation of the peptide, and its biomimetic folding, and the products **2**, **3**, and **4** were isolated by RP-HPLC. The typical yields of this last synthetic step were about 8% (0.5 mg) for **2**, 5.7% (0.27 mg) for **3**, and 1.7% (0.12 mg) for **4** (Figures S2–S4), which provided sufficient material for their biological characterization.

In summary, during the synthesis of peptide chain precursors and analogs, we encountered problems with the aspartimide formation in peptide **2a** and partly also with the oxidation of Met59 in peptide **2b**. We also found that S-sulfonate protection is not compatible with Cu^I-catalyzed click reaction and that simultaneous cleavage of six AcM-protecting groups in molecules of precursors with an access of I₂ was ineffective

and had to be rectified by the use of silver triflate, followed by the removal of Ag⁺ with a gaseous H₂S.

Functional Characterization of the analogs. Binding Affinities to the Receptors. The analogs **2–4**, and human insulin and IGF-1 as controls, were tested for their binding to human IGF-1R in membranes of mouse fibroblasts and to human IR-A in membranes of human IM-9 lymphocytes (Table 1, with the representative binding curves in Figures S5 and S6, respectively).

The two-chain analog **1** has significantly impaired binding for IGF-1R of only about 3.5% of the native IGF-1. Similar low binding (4.8%) was also detected for a single-chain analog **3**, in which Arg36 and Arg37 were replaced with β -azido-alanine and propargylglycine, respectively, connected by a triazole link, and with Met59Nle mutation. Interestingly, analog **2**, a "wild-type" homologue of **3** with preserved native Met59, showed markedly stronger 20.8% affinity than analog **3**.

The insertion of Arg36 and Arg39 in analog **4**, adjacent to the connecting triazole-containing 37–38 moiety, further enhanced IGF-1R binding affinity of this analog to almost 25% of the native IGF-1 binding.

Native human IGF-1 has only 1.0% binding affinity for IR-A compared to human insulin, but all analogs were even significantly weaker than human IGF-1 in binding to this receptor (Table 1).

Autophosphorylation of Receptors and Akt Activation. We also investigated the ability of analogs **1–4** to induce the autophosphorylation of IGF-1R (Figure 3A,C), and subsequent downstream phosphorylation of the Akt protein (pAkt) (Figure 3B,D), in response to their binding to this receptor. The potencies of the analogs were measured in IGF-1R-transfected 3T3 fibroblasts with high receptor population (Figure 3A,B) and, in parallel, in the non-transfected fibroblasts with a natural, that is, much lower, mouse IGF-1R population (Figure 3C,D). We also determined the ability of analogs **1–4** to stimulate the autophosphorylation of IR-A receptor in the IR-A-transfected fibroblasts (Figure 3E). The cells were stimulated for 10 min at 10⁻⁸ M concentration, and signals compared for all analogs (Figure 3A–E). Subsequently, the dose–response curves were measured for each analog (Figure 3F).

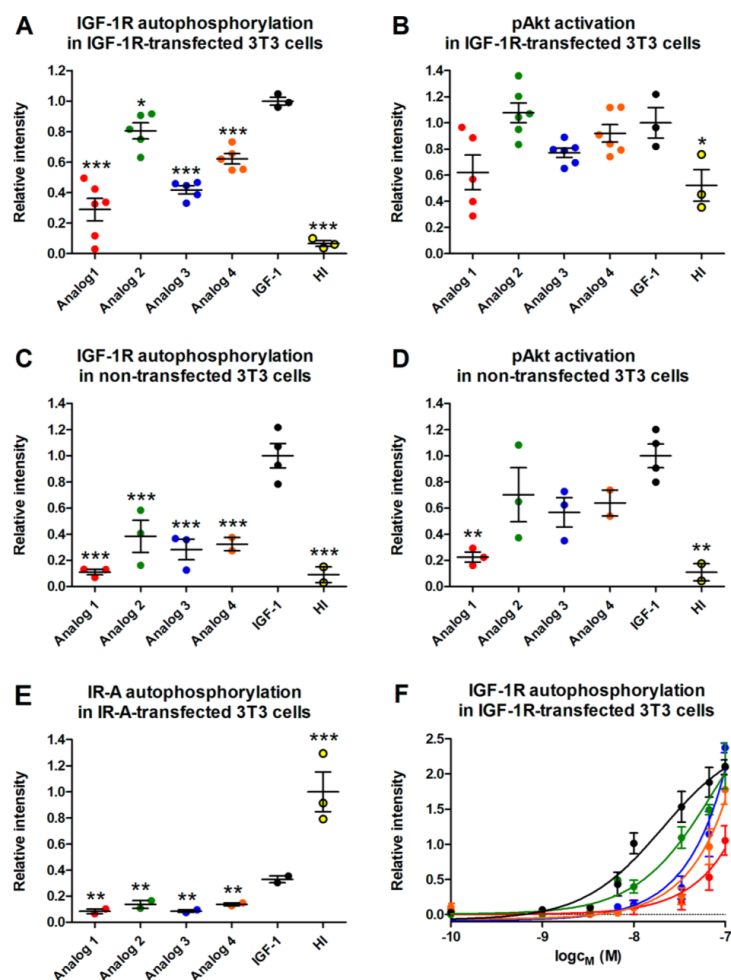


Figure 3. Receptor autophosphorylation and pAkt activation induced by analogs 1–4, human IGF-1, and human insulin (HI). In (A–E), the cells were stimulated for 10 min with proteins at 10^{-8} M concentration only. In (F), the dose responses for proteins at several different concentrations were determined. In all panels, the same color scheme was used: human IGF-1 in black, analog 1 in red, analog 2 in green, analog 3 in blue, analog 4 in orange, and human insulin (HI) in yellow. (A) IGF-1R autophosphorylation in IGF-1R-transfected 3T3 fibroblasts. (B) pAkt activation in IGF-1R-transfected 3T3 fibroblasts. (C) IGF-1R autophosphorylation in non-transfected 3T3 fibroblasts. (D) pAkt activation in non-transfected 3T3 fibroblasts. (E) IR-A autophosphorylation in IR-A-transfected 3T3 fibroblasts. In most cases, the scatter plots show means \pm SD of several independent measurements. However, in some cases where a low amount of material was available (for analog 4 in C–D), the plots show mean range calculated from two measurements. HI is human insulin. (F) Concentration-dependent effects of analogs on IGF-1R autophosphorylation in IGF-1R-transfected 3T3 fibroblasts. Asterisks indicate that phosphorylation of the receptor or Akt induced by the ligand differs significantly from that of IGF-1 (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

In general, the abilities of the analogs to stimulate autophosphorylation of IGF-1R and phosphorylation of intracellular Akt protein paralleled their IGF-1R binding affinities. The lowest IGF-1R autophosphorylation was observed for the two-chain analog 1 (4% binding) and for analog 3 with Met59Nle substitution (5% binding). However, interestingly, the analog 2 that lacks Arg36 and Arg37 activates IGF-1R somehow more strongly than analog 4, which has Arg36 and Arg39 (Figure 3A), despite their reversed binding affinities of 21% and 24%, respectively. The ability of these

analogues to activate/phosphorylate Akt protein follows this trend, although the differences between 2 and 4 are less significant here and similar to the native IGF-1 (Figure 3B). We compared the data measured for analogs 2 and 4, using the two-tailed t test to determine the statistical significance of the observed differences, and only the stimulation of autophosphorylation of IGF-1R in transfected fibroblasts appeared statistically significant ($p = 0.02$), while the difference in binding activities was negligible ($p = 0.82$). In general, it appears that the transfected cells (Figure 3A,B) respond more

sensitively to the stimulation with analogs than the non-transfected cells (Figure 3C,D), and this trend is visible especially in Akt activation (Figure 3B,D). It seems that higher sensitivity of the transfected cells alleviates the differences in binding affinities of analogs.

All analogs are also less potent than native IGF-1 in the stimulation of IR-A autophosphorylation in IR-A-transfected cells (Figure 3E).

Finally, the dose–response curves of IGF-1R autophosphorylation ability of analogs 1–4 and IGF-1 (Figure 3F) indicate that all analogs are weaker than IGF-1, where analog 2 is the strongest and analog 1 is the weakest, without a major difference between 3 and 4; all in general agreement with the previous results.

DISCUSSION

Chemistry. The low yield (0.6%) of the recombination reaction, that is, a biomimetic formation of disulfides, for analog 1 is similar to the rather poor yields of some insulin analogs prepared previously and is significantly lower than typical yields for native human insulin (~10%).⁴³ The poor folding ability of two separate chains of IGF-1 is also reflected in the complexity of the byproducts in the reaction mixture, where we identified, for example, a misfolded, inactive, but correct molecular mass, analog; the active analog 1 was found there as a rather minor peak (Figure S9). Misfolded and inactive IGF-1 molecules, specifically with “swapped” pairs of disulfides, that is, Cys6-Cys47 and Cys48-Cys52, instead of the correct Cys6-Cys48 and Cys47-Cys52, were identified previously in preparations of synthetic IGF-1 and its analogs.²⁰ However, the particularly poor yield of 1 clearly indicates the importance of a single-chain form of IGF-1 for its effective folding. Nevertheless, the amounts of analog 1 were sufficient for its biological characterization, especially with the recycling of chains 1a and 1b and repetition of the reaction.

We assumed that the single-chain IGF-1 analogs, with six free -SH groups, will oxidize themselves to a native conformation and that better folding properties of the single, rather than two-chain, IGF-1 analogs can be expected. However, as mentioned above, the 70-amino acid chain of IGF-1 is too long for an efficient solid-phase peptide synthesis. Therefore, we investigated here the possibility of connecting two IGF-1 fragments by Cu^I-catalyzed 1,3-dipolar cycloaddition of azides and alkynes, which would generate 1,2,3-triazole in the main-chain of IGF-1 analog. A similar approach was used by Valverde et al.²⁷ to connect three fragments of cysteine-protease protein inhibitor cystatin (98 amino acids). However, this did not require the formation of native disulfide bonds in this protein. As 1,4-disubstituted 1,2,3-triazoles are considered to be good mimics of the peptide bond,²⁸ it might be possible that their incorporation into the peptide chain in the IGF-1 C domain could be tolerated by IGF-1R. We were also encouraged here by our similar approach in the cross-linking of insulin B-chain with 1,2,3-triazole bridges, in which some highly potent IR analogs were obtained.⁵⁷

Our first attempts at the clicking of two IGF-1 fragments involved the use of IGF-1 fragment 1–36 with β -azido alanine at position 36 and the IGF-1 fragment 37–70 with the propargyl glycine at position 37; all cysteines in these peptides were protected as *S*-sulfonates, which assured their good solubility and better separation on the HPLC. However, all our experiments (not shown) for connecting these fragments by Cu^I-catalyzed click reaction were unsuccessful (Figure S1). We

concluded that the *S*-sulfonate groups are incompatible with the reaction, probably due to the *S*-sulfonate-copper chelation, which could lead to an undesired Cu-mediated folding of IGF-1 chains and steric blocking of the azido and alkyne groups, making them inaccessible for cycloaddition.

For this reason, we changed the Cys-protection strategy, synthesizing peptide chains with acetamidomethyl-protected cysteines, assuming that the Ac_m protection should not interfere with the clicking step, offering also several different alternative deprotections.⁶¹ Unfortunately, the first synthesis of 2a was hampered by a considerable amount of other, hardly separable, byproducts, identified as racemized aspartimides at positions Asp20-Arg21 (i.e., peptide 2ax, see Figures S2, S12, and S28). The appearance of Asp-Arg aspartimides is rather rare, but not unusual in peptide chemistry,⁶² as the Arg(Pbf)/Cys(Ac_m) protection strategy may promote this side reaction to a greater extent.⁶³ To test such a scenario in our chemical environment, we performed a model experiment (Supporting Information, pp 27–28) in which, with a model peptide 7, Phe-Val-Cys(Ac_m)-Gly-Asp(OtBu)-Arg(Pbf), we were able to generate and characterize (Figures S29, S30 and Scheme S1, Tables S1–S3) racemized aspartimides that were similar to the ones found in 2ax. Therefore, we used a Fmoc-Asp(OEpe) derivative,⁶⁴ which helped to eliminate this problem in 2a and 4a. Hence, it seems that the formation of Asp20-Arg21 aspartimide in 2ax was both sequence and side-chain protection-type dependent, as such byproducts were not observed in 1a, 2b, 3a, and 4b.

As mentioned above, the effective reactions and isolation of the products in the cycloaddition step required a reaction setup with a high excess of CuSO₄/ascorbate in water-*t*BuOH,^{56,57} but the addition of some click accelerators^{65,66} or reactions in organic solvents²⁵ did not lead to the desired products. It is possible that the participation of bulky copper chelating agents like TBTA or THPTA is not favorable for steric reasons, if large peptide precursors are used. The failure of CuBr or CuI/DIPEA strategy in DMF could be due to a need for an effective protein hydration that is better in water-*t*BuOH. Although all reactions performed here were ineffective if they were scaled up above ~2–4 mg per chain, some amount of the precursors could be recycled, and the reactions repeated. Both single-chain peptides without arginines in the proximity of the azido and alkyne precursor residues (2c and 3b) gave better yields than peptide 4c with the triazole-neighboring arginines. This could be due some steric shielding of the azido and alkyne moieties by the bulky arginine side chains.

Overall, we were able to prepare full-length IGF-1 chains on a preparative scale, still within reasonable yield ranges (7–18%). Our results underline the difficulty of click reactions where both precursors are large peptides, or proteins, compared to click reactions with small peptides/molecules, which usually proceed quickly and quantitatively.^{25,26}

The final synthetic steps, that is, the deprotection of cysteines and the formation of physiological S–S bridges, were crucial for the success of the synthesis. Interestingly, the use of iodine in acetic acid/HCl⁶¹ was not successful here, due to an incomplete cleavage of the Ac_m groups. One of the reasons for this could be the presence of six Ac_m groups accumulated in each molecule of the precursor, while only two Ac_m groups are usually simultaneously deprotected by I₂ for a directed formation of the disulfide.³⁰ For this reason, we used silver triflate, instead of I₂, for the cleavage of the Ac_m groups, without the purification of an intermediate product. Interest-

ingly, this three-step procedure was relatively effective and successful, providing 6–8% yields for analogs 2 and 3, which are comparable to our yields of syntheses of human insulin from its separate A and B chains (~10%). However, calculated for a single folding step,⁴³ they were also much better than the ~0.6% yields for the two-chain IGF-1 analog 1. The less effective (1.7%) folding of analog 4 may result from the presence of Arg36 and Arg39 in this analog and, consequently, from a longer C domain than in 2 and 3. It seems that the insertion of two additional Arg residues, rather than the presence of the triazole-containing connecting moiety, lowered the folding efficiency of this IGF-1 precursor. The folding of 2–4 was accompanied by the appearance of inactive misfolded “swapped” products (Figure S16), but to a lesser extent than in the case of 1 (Figure S9), certainly due to the single-chain character of precursors for 2–4.

The clicked polypeptides reported here are, to our knowledge, unique examples of the use of triazole-forming chemistry for the preparative production of large peptides/small proteins (relative molecular weights about 8000 Da), with a native pattern of disulfide bridges. The middle position of the connecting triazole bridge in the protein main chain further validates the results. Moreover, the only few examples of true proteins with main-chain triazoles that were reported^{27,67} did not contain disulfide bridges, which present additional challenges to such an approach.

Biology. Analog 1 retained some IGF-1R binding affinity, but its binding potency was severely impaired (3.5%) in comparison with the potency of the native IGF-1 (Table 1). This strongly suggests that the cleavage of Arg36–Arg37 peptide bond, if it occurs *in vivo*, is not a part of IGF-1 processing, but is rather a degradation-like step of this hormone, which leads to a much less active protein. Whether such cleavage represents some physiological attenuation of IGF-1 activity cannot be excluded, but it requires further, complex studies on this hormone, especially on *in vivo*-derived material.

The incorporation of the triazole bridge formed from β -azido-alanine at 36 and propargylglycine at 37 reduced the binding affinity of the analog 2 to about 21% of the binding affinity of the native IGF-1. However, this affinity drop is less significant than the effect of Ala36–Ala37 mutation of this hormone, which reduced its receptor binding to about 5%.⁴⁷ Therefore, it seems that the loss of Arg36–Arg37 in 2 is, somehow, compensated for by the triazole ring and/or the neighboring ethyl carboxymido (at 36) and ethyl amino (at 37) moieties. Some insight into the functional differences of the analogs 1–4 may be facilitated by a direct comparison of their connecting motifs shown in Figure 4. In 2, this motif is three atoms longer than the peptide bond of the native IGF-1, and the only difference between analog 2 and 3 is Met59Nle mutation in analog 3. This exchange reduced the IGF-1R affinity of the analog 3 to 5%, in agreement with the work of Shooter et al.,⁶⁸ where Met59Phe mutation resulted in the 17-fold decrease in IGF-1R binding. Therefore, analog 3 further highlights the importance of the sulfur atom in Met59 for receptor binding, which is not compensated for in 3 by the Nle alkyl chain. The highest IGF-1R binding affinity in the series from Table 1 was found for analog 4 (almost 25% of native IGF-1). However, statistical analysis (see above) showed that the difference in binding affinities of 4 and 2 is not significant, hence, both analogs are rather similar in this aspect. Although analog 4 retains arginines in the C-loop, they are at sites 36 and 39, not at their native positions 36–37 (Figures 2 and 4).

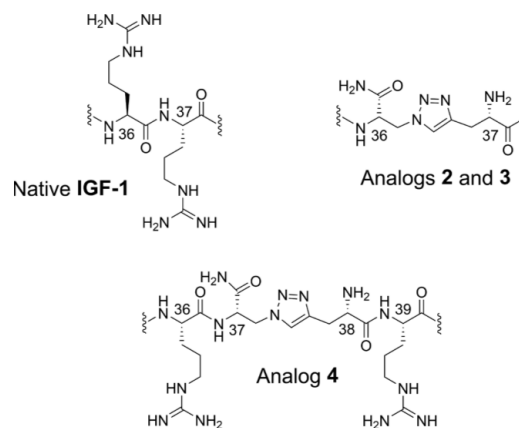
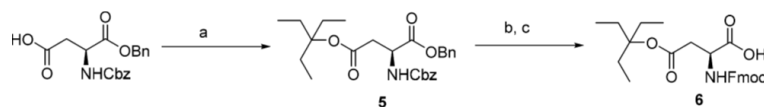


Figure 4. Comparison of the connecting motifs at positions 36–37 or 37–38 in analogs 2–4 with the native human IGF-1. C α atoms are marked by residue numbers.

Therefore, the C α atoms of the arginines in 4 are separated by eight extra backbone atoms plus triazole ring and carboxamide and amino moiety. Interestingly, this extension did not reduce the binding affinity of this analog in comparison with analog 2, indicating a relative tolerance of IGF-1R for the modifications in this part of the C domain. However, it is also possible that at least one of the Arg36/Arg39 in 4 (e.g., Arg36) occupies a similar position on the receptor as one of the natives Arg36–37. This result opens up new possibilities for the use of different side chains, and motifs, at positions adjacent to the triazole ring, that is, 36 and 37 in 2 and 37 and 38 in 4. However, such approaches depend on the development of available functionalized azides and alkynes suitable for the solid-phase peptide synthesis.

All analogs 1–4 have reduced binding affinities for the isoform A of the insulin receptor (Table 1) compared to native IGF-1. However, it seems that this decrease (21–43% of native IGF-1) is lower than for the IGF-1R (3–25% of native IGF-1), indicating that the C domain of IGF-1 is less important for IR-A binding than for an effective hormone complex formation with the IGF-1R.

The abilities of analogs 1–4 to induce the autophosphorylation of IGF-1R and IR-A and to activate the downstream Akt generally reflected their binding affinities for these receptors. This trend is especially visible in non-transfected fibroblasts, with a natural population of IGF-1R (Figure 3C,D). It seems that the high non-natural concentration of IGF-1R in IGF-1R-transfected fibroblasts is behind their higher sensitivity to the stimulation by the analogs, rather than differences in binding affinities of human IGF-1 for rat and human IGF-1R; this levels up the differences in binding affinities of analogs. The reason for this phenomenon could be in the “disproportional” populations/distributions of receptors and intracellular signaling proteins. In this respect, the native fibroblasts, with naturally lower population of IGF-1R, seem to be a better model. The only apparent exception here could be observed in a higher receptor stimulating activity of analog 2 than of analog 4 despite slightly reverse binding affinities. This “discrepancy” is interesting, but should be treated with caution, due to the low statistical significance of the differences between biological properties of analogs 2 and 4 mentioned above. Nevertheless,

Scheme 1^a

^aReagents, conditions, and yields (a) DIC, DMAP, 3-ethyl-3-pentanol, DCM 0 °C 1 h, then rt 48 h (91%); (b) 10% Pd/C, H₂, 10 psi, methanol 12 h; (c) Fmoc-OSu, NaHCO₃, water and dioxane, 0 °C 1 h, then rt overnight (65% over two steps).

we observed a similar phenomenon in the case of insulin-IGF hybrids, that is, human insulin molecules with the B chain prolonged for C domain amino acids of human IGF-2.⁵⁵ However, it is rather difficult to explain such a discrepancy without a structure of the analogs in a complex with the receptor, and even crystal/NMR structures of the analogs would provide only very limited clues. It is possible that only a complex with the so-called site 2 of IGF-1 receptor would shed light on the slightly differential activation/binding observed here, and only the structures of insulin and IGF-1 complexed with site 1 of the IR are currently available.^{69,70} Nevertheless, as it was observed that differential binding kinetics of insulin-like analogs, that is, slower or faster k_{on} or k_{off} rates, can influence their metabolic or mitogenic biological responses,⁷¹ this could be the case of the apparent differences between analogs 2 and 4.

Finally, analogs 1–4 do not appear to possess the properties of IGF-1R antagonists or partial antagonists, as was observed for the IGF-1 Glu36-Glu37 mutant.^{49,50} It is possible then that the antagonism observed for the Glu36-Glu37 IGF-1 analog results from a reversed, negatively charged glutamic acid's side chains, instead of the arginine basic guanidine groups, and that the replacement of the Arg36-Arg37 pair, by moieties shown in Figure 4, is not sufficient to elicit a similar antagonistic effect.

CONCLUSIONS

In summary, we developed a new method for the total chemical synthesis of human IGF-1 analogs, which entails a solid-phase synthesis of two IGF-1 precursor chains, and their ligation by Cu^I-catalyzed azide–alkyne cycloaddition chemistry. This relatively straightforward methodology presents a viable synthetic alternative to the previously established IGF synthetic strategies, which employed native ligation. The connection of two IGF-1 precursor chains by triazole-containing moieties was relatively well tolerated by IGF-1R and IR-A, both in terms of receptor binding and activation, and can thus represent a convenient and relatively straightforward synthetic strategy for the design of new analogs of this important human hormone with non-standard protein modifications. We also found that the disruption of the connectivity of the IGF-1 Arg36-Arg37 peptide bond deactivates this protein, hence it is unlikely that such cleavage represents a maturation step of this hormone in vivo.

EXPERIMENTAL SECTION

Unless otherwise stated, reagents and materials were obtained from commercial suppliers (Sigma-Aldrich-Fluka, Merck) and used without purification.

Solid-Phase Peptide Synthesis. Peptide chains were prepared by manual solid-phase synthesis in plastic syringes, equipped either with a Teflon frit or with ABI433A peptide synthesizer. Standard Fmoc amino acids (Iris Biotech, GmbH, Germany) were used, and couplings were done with HBTU/DIPEA or DIC/HOBT reagents in DMF. Usually two couplings (60 min) with Fmoc-amino acids (3–5 equiv

for manual synthesis and 10 equiv for automated synthesis) were done. In manual synthesis, the efficiency of the couplings was monitored by the Kaiser test, and the amount of cleaved Fmoc group (20% piperidine in DMF for 5 and 20 min) was determined spectrophotometrically (extinction coefficient 7040 M⁻¹ cm⁻¹ at 301 nm). Fmoc-Asp(OEpe)-OH was prepared as described in the Supporting Information and used for the syntheses of 2a and 4a. Fmoc-β-azido-(S)-Ala-OH was prepared as described by Pícha et al.⁷²

The peptide chains with the C-terminal β-azido-(S)-Ala carboxamide, 2a and 4a, were synthesized on Rink amide AM LL resin or Nova PEG Rink amide (Merck-Novabiochem), respectively. Fmoc-(S)-propargylglycine-OH was purchased from Sigma-Aldrich-Fluka. The peptide chains 1a, 1b, 2b, 3a, and 4b with a free carboxy terminus were synthesized on Wang-ChemMatrix LL resins, preloaded with Fmoc-Ala or Fmoc-Arg(Pbf) (PCAS BioMatrix Inc.). The peptides were cleaved with 91.5% TFA, 2.3% H₂O, 2.3% thioanisole, 2.3% phenol, 1.15% DODT, and 0.45% TIS for 2–4 h. Subsequently, the crude peptides were precipitated from the cold diethyl ether (in at least 10-fold excess), centrifuged, and dried. The precipitates of crude peptides 1a and 1b were immediately converted to S-sulfonates and subsequently desalted, as previously described in detail.^{42,43} Finally, all peptide chains were purified by RP-HPLC.

Peptide and Compound Characterization. The precursor peptides (1a, 1b, 2a, 2b, 2c, 3a, 3b, 4a, 4b, and 4c), and the resulting IGF-1 analogs 1–4 were purified and the purity of compounds analyzed by RP-HPLC. HRMS spectra of all compounds, peptides, and analogs were obtained on a FTMS mass spectrometer LTQ-orbitrap XL (Thermo Fisher, Bremen, Germany) in electrospray ionization mode. The purity of target compounds 1–4 was ≥95% (RP-HPLC, HR-MS). The identity and purity of analogs 1–4 were furthermore verified by N-terminal sequencing. Analytical data for all peptide precursors, IGF-1 analogs 1–4, and conditions for NMR measurements of 2a and model compounds 7 are provided in the Figures S7–S30, Tables S1–S3, and Scheme S1. Melting points of compounds 5 and 6 were determined on a Boetius block and are uncorrected. ¹H and ¹³C-NMR spectra of compounds 5 and 6 were measured on a Bruker AVANCE-600 spectrometer (¹H at 600.13 MHz, ¹³C at 150.9 MHz) in CDCl₃, DMSO-*d*₆, CD₃OD, or D₂O solution at 300 K. The 2D-H,H-COSY, 2D-H,C-HSQC and 2D-H,C-HMBC spectra were recorded and used for the structural assignment of proton and carbon signals. IR spectra of compounds 5 and 6 were recorded on Bruker IFS 55 Equinox apparatus.

Peptide 1a. The yield (RP-HPLC) was 4.3% (27 mg). HRMS-ESI: Monoisotopic M_r calcd for C₁₆₉H₂₅₃N₄₅O₆₀S₄ 4000.701, found 4000.677.

Peptide 1b. The yield (RP-HPLC) was 12.2% (51 mg). HRMS-ESI: Monoisotopic M_r calcd for C₁₆₂H₂₆₇N₄₉O₆₀S₅ 4146.683, found 4146.665.

Analog 1. Recombination of chains 1a (17 mg, 4.27 μmol) and 1b (33 mg, 7.95 μmol) was done according to the procedure previously described in detail.^{42,43} The typical yield of the analog 1 (relative to chain 1a as the limiting component of the reaction) after RP-HPLC purification was about 0.6% (0.2 mg). HRMS-ESI: Monoisotopic M_r calcd for C₃₃₁H₅₁₄N₉₄O₁₀₂S₇ 7661.597, found 7661.617.

Peptide 2a. When Fmoc-Asp(OtBu)-OH was used, the yield (RP-HPLC) was 6.2% (22 mg). HRMS-ESI: M_r calcd for C₁₇₂H₂₅₆N₄₈O₅₅S₂ 3937.815, found 3937.818. The aspartamide derivative of 2a (with an aspartamide bond at Asp20-Arg21, that is, peptide 2ax, see in the Supporting Information) was purified (Figure

S28) as well, and the yield was 8.6% (30 mg). HRMS-ESI M_r calcd for $C_{172}H_{254}N_{48}O_{54}S_2$ 3919.805, found 3919.809.

If Fmoc-Asp(OEpe)-OH (**6**) was used for the synthesis, the yield (RP-HPLC) was 10.1% (120 mg). HRMS-ESI: M_r calcd for $C_{172}H_{256}N_{48}O_{55}S_2$ 3937.815, found 3937.819. Undesired aspartimide byproduct was not found in this case.

Peptide 2b. The yield (RP-HPLC) was 18.5% (75 mg). HRMS-ESI M_r calcd for $C_{173}H_{280}N_{50}O_{52}S_3$ 4049.941, found 4049.945.

Peptide 2c. In general, the click reactions were performed according to the protocol previously published by Vikova et al.,⁵⁷ but with a few modifications. Briefly, the typical reaction mixture contained the linear peptide **2a** (4 mg, 1.02 μ mol) and **2b** (2 mg, 0.49 μ mol), which were dissolved in 4.5 mL of a *tert*-butanol:water mixture (2:1, v/v). Next, the ascorbic acid (5.5 μ mol) and $CuSO_4 \cdot 5 H_2O$ (4.80 μ mol) were added in a minimum volume of water (the mixture was freshly prepared). The reaction mixture proceeded with stirring at 40 °C overnight. The precipitated product was removed from the reaction mixture by centrifugation and was purified using RP-HPLC (C4 column). The typical yield (RP-HPLC, relative to the chain **2b** as the limiting component of the reaction) was about 16% (0.6 mg). HRMS-ESI M_r calcd for $C_{345}H_{536}N_{98}O_{107}S_7$ 7987.756, found 7987.768.

Analogue 2. Clicked peptide **2c** (6.6 mg, 0.83 μ mol) was dissolved in 1 mL of TEA with 250 mM silver triflate and 25 mM anisole and allowed to react for 1 h in an ice bath. The deprotected peptide was removed from the solution by precipitation with cold diethyl ether. The crude peptide was dissolved (60 μ M) in degassed 0.1 M Gly/NaOH buffer (pH 10.5), and then the reaction mixture was treated with gaseous H_2S^{58-60} to remove Ag^+ ions from cysteine salts. The precipitated Ag_2S was removed by centrifugation. The supernatant, containing the peptide with free SH groups, was diluted to 30 μ M concentration (by the same buffer) and purged with argon to remove the excess hydrogen sulfide. The reaction mixture was briefly purged with air and allowed to oxidize in an opened flask without stirring for 3 days at 4 °C. The crude target compound **2** was desalted on a C4-cartridge (Macherey-Nagel) and purified using RP-HPLC (C4 column). The yield (RP-HPLC) was 8% (0.5 mg). HRMS-ESI M_r calcd for $C_{327}H_{500}N_{92}O_{101}S_7$ 7555.486, found 7555.502.

Peptide 3a. The yield (RP-HPLC) was 15% (46 mg). HRMS-ESI M_r calcd for $C_{174}H_{282}N_{50}O_{52}S_4$ 4031.984, found 4031.988.

Peptide 3b. The compound was prepared from **3a** (2 mg, 0.5 μ mol) and **2a** (2 mg, 0.51 μ mol) according to the procedure described for **2c**. The typical yield (RP-HPLC) was about 18.5% (0.74 mg). HRMS-ESI M_r calcd for $C_{346}H_{538}N_{98}O_{107}S_6$ 7969.799, found 7969.801.

Analogue 3. Analogue **3** was prepared from **3b** (5 mg, 0.63 μ mol) by the method described for **2**. The yield (RP-HPLC) was 5.7% (0.27 mg). HRMS-ESI M_r calcd for $C_{328}H_{502}N_{92}O_{101}S_6$ 7537.530, found 7537.551.

Peptide 4a. The yield (RP-HPLC) was 13.8% (170 mg). HRMS-ESI M_r calcd for $C_{178}H_{268}N_{52}O_{56}S_2$ for 4093.916, found 4093.923.

Peptide 4b. The yield (RP-HPLC) was 16.6% (140 mg; 33 μ mol). HRMS-ESI M_r calcd for $C_{179}H_{292}N_{54}O_{53}S_5$ 4206.042, found 4206.047.

Peptide 4c. Peptide **4c** was prepared from peptides **4a** (4 mg, 0.98 μ mol) and **4b** (2 mg, 0.48 μ mol) by the method described for **2c**. The typical yield (RP-HPLC, relative to the chain **4b** as the limiting component of the reaction) was 7.3% (0.29 mg). MS-ESI M_r calcd for $C_{357}H_{560}N_{106}O_{109}S_7$ 8299.958, found 8299.987.

Analogue 4. Analogue **4** was prepared from **4c** (7.5 mg, 0.9 μ mol) by the method described for analogue **2**. The yield (RP-HPLC) was 1.7% (0.12 mg). HRMS-ESI M_r calcd for $C_{339}H_{524}N_{100}O_{103}S_7$ 7867.688, found 7867.709.

(2S)-N-Benzoyloxycarbonylaspartic Acid 1-Benzyl-4-((3-ethyl)pent-3-yl)ester 5. Intermediate **5** was prepared using a slightly modified method described previously in a patent.⁶⁴ Briefly, DIC (870 μ L; 5.63 mmol) was added under stirring and ice-cooling to a solution of Cbz-Asp-OBn (5 g; 14 mmol) and DMAP (0.34 g; 2.8 mmol), and 3-ethyl-3-pentanol (4.8 g; 42 mmol) in 25 mL DCM was added DIC. The cooling was finished, and the mixture was allowed to react at rt 16 h. Addition of DIC (870 μ L) was repeated twice, and the total reaction time was 48 h. The excess of DIC was eliminated by quenching with glacial acetic acid, and the volatile material was evaporated in vacuo to give a yellow residue, which was purified by flash chromatography on

silica gel using a linear gradient of ethyl acetate in petroleum ether. Yield 5.8 g (91%). Colorless oil. R_f = 0.74 (20% ethyl acetate/toluene). $[\alpha]_D^{20}$ = +14° (c = 0.669; $CHCl_3$). 1H NMR (500 MHz, $CDCl_3$): 0.70 (9H, t, J = 7.5, 3 \times CH_3), 1.71 (6H, q, J = 7.5, 3 \times CH_2), 2.96 (1H, dd, J = 17.2 and 4.5, CO- CH_2 Hb-), 2.77 (1H, dd, J = 17.2 and 4.5, CO- CH_2 Hb-), 4.56 (1H, dt, J = 8.8, 4.5 and 4.5, >CH-N), 5.05 (1H, d, J = 12.2, O- CH_2 Hb-), 5.08 (1H, d, J = 12.2, O- CH_2 Hb-), 5.10 (1H, d, J = 12.4, O- CH_2 Hb-), 5.14 (1H, d, J = 12.4, O- CH_2 Hb-), 5.77 (1H, d, J = 8.7, -NH-CO), 7.25–7.30 (10H, m, 2 \times C_6H_5); ^{13}C NMR (125.7 MHz, $CDCl_3$): 7.53 (3 \times CH_3); 26.63 (3 \times CH_2); 37.27 (CH_2); 50.53 (CH-N); 66.95 and 67.40 (2 \times CH_2 -O), 90.06 (C-O), 127.95(2), 128.08, 128.17(2), 128.32, 128.45(2) and 128.49(2) (10 \times Ar =CH-), 135.18 and 136.18 (2 \times Ar > C=), 155.98 (N-CO), 169.61 and 170.75 (2 \times O-CO). IR (CCl_4) ν_{max} cm^{-1} 3433 m, 3353 m (NH); 2972 s, 2883 m, 1380 s (CH_3); 2943 s (CH_2); 1751 vs + br (C=O) esters; 1727 vs (C=O) carbamate; 3091 w, 3066 w, 3034 w, 1500 s, 1456 s, 1175 s, 1028 s, 1003 s, 737 s, 697 m (ring). HRMS (ESI) calcd for $C_{26}H_{33}O_6NNa$ [$M + Na$]⁺ 478.220, found: 478.220.

The synthesis of Fmoc-Asp(OEpe)-OH **6** is summarized in Scheme 1.

(2S)-N-((9H-Fluoren-9-ylmethoxy)carbonyl)aspartic acid 4-((3-ethyl)pent-3-yl)ester 6. Intermediate **5** (6.9 g; 15.1 mmol) was charged to a nitrogen flushed glass pressure bottle, dissolved in 80 mL of methanol, and then 500 mg of 10% Pd/C was added. The mixture was vigorously stirred and allowed to react under the atmosphere of hydrogen (15 psi) at RT for 12 h. TLC analysis (50% ethyl acetate/toluene) revealed that the starting compound had completely disappeared. The catalyst was filtered off through Celite, and the filter cake was washed with 100 mL of methanol. The filtrate was evaporated in vacuo to give 3.5 g of light brown residue, which was immediately dissolved in 50 mL of aqueous solution of $NaHCO_3$ (2.5 g; 30.2 mmol). The flask was placed in an ice bath, and Fmoc-OSu (5.1 g; 15.1 mmol) in 50 mL of dioxane was added dropwise under stirring. After the addition of the total amount of Fmoc-OSu, stirring continued for 1 h at 0 °C and then overnight at room temperature. Thereafter, 10% aqueous solution of citric acid was added to the reaction mixture until pH ~3–4 was reached, followed by 200 mL water. The solution was transferred to a separator funnel and extracted with 4 \times 100 mL of ethyl acetate. Combined organic layers were washed consecutively with 2 \times 50 mL water, 2 \times 50 mL brine, and dried over anhydrous Na_2SO_4 . The filtrate was evaporated to afford a dark yellow oil, which was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in toluene. Yield 4.5 g (65% over two steps). Bright yellow semisolid. R_f = 0.67 (ethyl acetate-acetone-ethanol-water 6/1/1/0.5). $[\alpha]_D^{20}$ = -8.6° (c = 0.326; DMF). 1H NMR (600 MHz, DMSO): 0.73 (9H, t, J = 7.5, 3 \times CH_3), 1.72 (6H, q, J = 7.5, 3 \times CH_2), 2.53 (1H, dd, J = 15.9 and 9.1, CO- CH_2 Hb-), 2.78 (1H, dd, J = 15.9 and 4.6, CO- CH_2 Hb-), 4.20 (t, J = 6.5, >CH-), 4.25 (3H, m, >CH-N and - CH_2 -O), 7.38 (1H, br, -NH-CO), 7.31 (2H, m), 7.41 (2H, m), 7.69 (2H, m) and 7.88 (2H, m) (2 \times C_6H_4); ^{13}C NMR (150.9 MHz, DMSO): 7.60 (3 \times CH_3); 26.50 (3 \times CH_2); 37.27 (CH_2); 46.83 (>CH-); 66.81 (CH_2 -O), 87.83 (C-O), 120.31(2), 125.44(2), 127.26(2) and 124.83(2) (8 \times Ar =CH-), 140.91(2), 144.02 and 144.02 (4 \times Ar > C=), 155.98 (N-CO), 169.71 (O-CO). IR (CCl_4) ν_{max} cm^{-1} 3437 w, 3322 w (NH); 2971 m, 2887 m, 1380 m, 1370 m (CH_3); 2945 m, 2856 m (CH_2); 1727 vs+vbr (C=O) all carbonyles; 1505 s (amide II); 3068 w, 3042 w, 3021 w, 1600 m, 1478 m, 1452 m, 678 s (ring). HRMS (ESI) calcd for $C_{26}H_{31}O_6NNa$ [$M + Na$]⁺ 476.204, found: 476.204.

Cell Cultures. Human IM-9 lymphocytes and NIH/3T3 mouse fibroblasts were obtained from ATCC. Mouse fibroblasts derived from IGF-1R knockout mice⁷⁵ and stably transfected with either human IGF-1 receptor (IGF-1R), or human insulin receptor isoform A (IR-A), were kindly provided by Professor A. Belfiore (University of Magna Grecia, Catanzaro, Italy) and Professor R. Baserga (Thomas Jefferson University, Philadelphia, Pennsylvania, USA). IM-9 cells were grown in RPMI 1640 media with 10% fetal bovine serum. IGF-1R and IR-A cells were grown in DMEM media with 5 mM glucose with 10% fetal bovine serum and 0.3 μ g/mL puromycin. NIH/3T3 cells were grown in DMEM media with 10% bovine serum albumin. All media

were further supplemented with 2 mM L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, and the cells were grown in humidified air with 5% CO₂ at 37 °C.

Receptor-Binding Studies. *Human IM-9 Lymphocytes (Human IR-A Isoform).* Receptor-binding studies with the insulin receptor in membranes of human IM-9 lymphocytes (containing only human IR-A isoform) were performed and K_d values determined as described previously.^{12,43} Binding data were analyzed using the Excel algorithms specifically developed for the IM-9 cells system in the laboratory of Prof. Pierre De Meyts (A. V. Groth and R. M. Shymko, Hagedorn Research Institute, Denmark, a kind gift from P. De Meyts), using the method of a non-linear regression, and a one-site fitting program, taking into account the potential depletion of the free ligand. Each binding curve was determined in duplicate, and the final dissociation constant (K_d) of an analog was calculated from at least three ($n \geq 3$) binding curves (K_d values), determined independently. The dissociation constant of human ¹²⁵I-insulin (PerkinElmer) was set to 0.3 nM.

Mouse Embryonic Fibroblasts (Human IGF-1R). Receptor binding studies with the human IGF-1 receptor in membranes of mouse embryonic fibroblasts derived from IGF-1R knockout mice and transfected with human IGF-1R were performed as described previously.^{12,43} Binding data were analyzed, and the dissociation constant (K_d) was determined with GraphPad Prism 5 software, using a non-linear regression method and a one-site fitting program, taking into account potential depletion of free ligand. Each binding curve was determined in duplicate, and the final dissociation constant (K_d) of an analogue was calculated from at least three ($n \geq 3$) binding curves (K_d values) determined independently. The dissociation constant of human ¹²⁵I-IGF-1 (PerkinElmer) was set to 0.2 nM. Here we should note that the use of bovine serum albumin (e.g., Sigma-Aldrich A6003) void of "IGF-binding-like" proteins, which interfere with these binding assays, is essential for the preparation of the binding buffer.⁷⁴

The significance of the changes in binding affinities of the analogs, related to the binding of IGF-1, was calculated using a two-tailed *t* test.

Receptor Phosphorylation Assay. *Cell Stimulation and Detection of Receptor phosphorylation.* Both were performed as previously described.⁵⁵ In brief, mouse fibroblasts (NIH/3T3, IR-A and IGF-1R) were seeded in 24-well plates (Schoeller) and incubated for 24 h. Cells were afterward starved for 4 h in serum-free media. The cells were stimulated with 10 nM concentrations of the ligands for 10 min. Proteins were routinely analyzed, using immunoblotting and horseradish peroxidase-labeled secondary antibodies (Sigma-Aldrich). The membranes were probed with antiphospho-IGF-1Rβ (Tyr1135/1136)/IRβ (Tyr1150/1151) and antiphospho-Akt (Thr308) (C31ESE) (Cell Signaling Technology). The blots were developed using the SuperSignal West Femto maximum sensitivity substrate (Pierce) and analyzed using the ChemiDoc MP Imaging System (Bio-Rad). Each experiment was repeated four times. The data were expressed as contribution of phosphorylation relatively to the human insulin (IR-A) respective IGF-1 (NIH/3T3 and IGF-1R) signal. The significance of the changes in stimulation of phosphorylation of receptors and related Akt-stimulation by the IGF-1 was calculated using one-way analysis of variance.

Ligand-Dose Response IGF-1R Autophosphorylation Levels. The levels were determined using In-Cell Western assay⁷⁵ adapted for chemiluminescence. The IGF-1R cells were plated at 20 000 cells/well in white 96-well Brand plates cell grade (Brand GMBH, Germany) and incubated for 24 h. Cells were afterward starved for 4 h in serum-free media and stimulated with dilutions of ligands (0–100 nM) for 10 min. After incubation, the medium was discarded, and the cells were fixed in 3.75% freshly prepared formaldehyde for 20 min. Cells were permeabilized with 0.1% Triton-X-100 in PBS for 5 min and blocked with 5% BSA in T-TBS for 1 h. Plates were incubated with antiphospho-IGF-1Rβ (Tyr1135/1136)/IRβ (Tyr1150/1151) overnight at 4 °C. Then plates were thoroughly washed with TBS, incubated with peroxidase-labeled antirabbit secondary antibody (Sigma) for 1 h at room temperature, and washed again. SuperSignal West Femto maximum sensitivity substrate was added to each well, and chemiluminescence was detected using the ChemiDoc MP Imaging System after 5 min. Data were subtracted from background

values and expressed as contribution of phosphorylation relatively to the 10 nM IGF-1 signal. Control wells and wells stimulated with 10 nM IGF-1 were in tetraplicates on each plate. Experiments were repeated at least four times. Log(agonist) vs response curve fitting of data was carried out with GraphPad Prism 5 software.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.7b01331.

Figures S1–S4: Synthetic pathways for analogs 1–4. Figures S5 and S6: binding curves of compounds with the receptors. Figures S7–S10, S12–S17, S19–S21, S23–S26: HPLC analyses of compounds. Figures S11, S18, S22 and S27: HRMS-ESI spectra of target compounds 1–4. Characterization of the aspartimide derivatives of peptides 2a and 7: Figures S28–S30, Tables S1–S3, Scheme S1 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +420 220183 441. E-mail: jiracek@uochb.cas.cz.

ORCID

Andrzej M. Brzozowski: 0000-0001-7426-8948

Jiří Jiráček: 0000-0003-3848-2773

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Medical Research Council Grant MR/K000179/1 (to A.M.B.). This work was also supported by the Research Project of the Czech Academy of Sciences RVO:61388963 (for the Institute of Organic Chemistry and Biochemistry).

■ ABBREVIATIONS USED

Acm, acetamidomethyl; c_m , molar concentration; DIC, *N,N'*-diisopropylcarbodiimide; DILP, *Drosophila* insulin-like peptide; DIPEA, *N,N*-diisopropylethylamine; DMF, dimethylformamide; DODT, 2,2'-(ethylenedioxy)diethanethiol; Epe, ethylpentyl; INSL, insulin-like peptide insulin receptor isoform A; IGF, insulin-like growth factor; IR-B, insulin receptor isoform B; IGF-1R, receptor for IGF-1; HBTU, *N,N,N',N'*-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate; HOBt, hydroxybenzotriazole; M_r , relative molecular weight; HRMS-ESI, high-resolution mass spectrometry electrospray ionization; RP-HPLC, reverse-phase high-performance liquid chromatography; TBTA, tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine; TFA, trifluoroacetic acid; THPTA, tris(3-hydroxypropyltriazolylmethyl)amine; TIS, triisopropylsilane; Trt, trityl

■ REFERENCES

- (1) Taniguchi, C. M.; Emanuelli, B.; Kahn, C. R. Critical nodes in signalling pathways: insights into insulin action. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 85–96.
- (2) Cohen, P. Timeline - the twentieth century struggle to decipher insulin signalling. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 867–873.
- (3) Denley, A.; Cosgrove, L. J.; Booker, G. W.; Wallace, J. C.; Forbes, B. E. Molecular interactions of the IGF system. *Cytokine Growth Factor Rev.* **2005**, *16*, 421–439.

- (4) Blüher, M.; Kahn, B. B.; Kahn, C. R. Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science* **2003**, *299*, 572–574.
- (5) Holzenberger, M.; Dupont, J.; Ducos, B.; Leneuve, P.; Geloën, A.; Even, P. C.; Cervera, P.; Le Bouc, Y. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* **2003**, *421*, 182–187.
- (6) Gallagher, E. J.; LeRoith, D. Minireview: IGF, insulin, and cancer. *Endocrinology* **2011**, *152*, 2546–2551.
- (7) Belfiore, A.; Frasca, F.; Pandini, G.; Sciacca, L.; Vigneri, R. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr. Rev.* **2009**, *30*, 586–623.
- (8) Mayer, J. P.; Zhang, F.; DiMarchi, R. D. Insulin structure and function. *Biopolymers* **2007**, *88*, 687–713.
- (9) Zaykov, A. N.; Mayer, J. P.; DiMarchi, R. D. Pursuit of a perfect insulin. *Nat. Rev. Drug Discovery* **2016**, *15*, 425–439.
- (10) Vajo, Z.; Fawcett, J.; Duckworth, W. C. Recombinant DNA technology in the treatment of diabetes: insulin analogs. *Endocr. Rev.* **2001**, *22*, 706–717.
- (11) Bright, G. M. Recombinant IGF-I: Past, present and future. *Growth Horm. IGF Res.* **2016**, *28*, 62–65.
- (12) Hexnerova, R.; Krizkova, K.; Fabry, M.; Sieglowa, I.; Kedrova, K.; Collinsova, M.; Ullrichova, P.; Srb, P.; Williams, C.; Crump, M. P.; Tosner, Z.; Jiracek, J.; Veverka, V.; Zakova, L. Probing receptor specificity by sampling the conformational space of the insulin-like growth factor II C-domain. *J. Biol. Chem.* **2016**, *291*, 21234–21245.
- (13) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.
- (14) Li, C. H.; Yamashiro, D.; Gospodarowicz, D.; Kaplan, S. L.; Vanvliet, G. Total synthesis of insulin-like growth factor-I (somatomedin-C). *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*, 2216–2220.
- (15) Iwai, M.; Kobayashi, M.; Tamura, K.; Ishii, Y.; Yamada, H.; Niwa, M. Direct identification of disulfide bond linkages in human insulin-like growth factor-I (IGF-I) by chemical synthesis. *J. Biochem.* **1989**, *106*, 949–951.
- (16) Bagley, C. J.; Otteson, K. M.; May, B. L.; Mccurdy, S. N.; Pierce, L.; Ballard, F. J.; Wallace, J. C. Synthesis of insulin-like growth factor-I using N-methyl pyrrolidinone as the coupling solvent and trifluoromethane sulfonic-acid cleavage from the resin. *Int. J. Pept. Protein Res.* **1990**, *36*, 356–361.
- (17) Luisier, S.; Avital-Shmilovici, M.; Weiss, M. A.; Kent, S. B. H. Total chemical synthesis of human proinsulin. *Chem. Commun.* **2009**, *46*, 8177–8179.
- (18) Sohma, Y.; Hua, Q. X.; Whittaker, J.; Weiss, M. A.; Kent, S. B. H. Design and folding of [Glu(A4)(O(beta)Thr(B30))]insulin ("ester insulin"): a minimal proinsulin surrogate that can be chemically converted into human insulin. *Angew. Chem., Int. Ed.* **2010**, *49*, 5489–5493.
- (19) Avital-Shmilovici, M.; Mandal, K.; Gates, Z. P.; Phillips, N. B.; Weiss, M. A.; Kent, S. B. H. Fully convergent chemical synthesis of ester insulin: determination of the high resolution X-ray structure by racemic protein crystallography. *J. Am. Chem. Soc.* **2013**, *135*, 3173–3185.
- (20) Sohma, Y.; Pentelute, B. L.; Whittaker, J.; Hua, Q. X.; Whittaker, L. J.; Weiss, M. A.; Kent, S. B. H. Comparative properties of insulin-like growth factor I (IGF-1) and [Gly7D-Ala]IGF-1 prepared by total chemical synthesis. *Angew. Chem., Int. Ed.* **2008**, *47*, 1102–1106.
- (21) Cottam, J. M.; Scanlon, D. B.; Karas, J. A.; Calabrese, A. N.; Pukala, T. L.; Forbes, B. E.; Wallace, J. C.; Abell, A. D. Chemical synthesis of a fluorescent IGF-II analogue. *Int. J. Pept. Res. Ther.* **2013**, *19*, 61–69.
- (22) Tornøe, C. W.; Meldal, M. Peptidotriazoles: Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions on Solid-Phase. In *Proceedings of the 2nd International and the 17th American Peptide Symposium*; Lebl, M., Houghten, R. A., Eds.; 2001; pp 263–264.
- (23) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.
- (24) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–3064.
- (25) Meldal, M.; Tornøe, C. W. Cu-catalyzed azide-alkyne cycloaddition. *Chem. Rev.* **2008**, *108*, 2952–3015.
- (26) Pedersen, D. S.; Abell, A. 1,2,3-triazoles in peptidomimetic chemistry. *Eur. J. Org. Chem.* **2011**, *2011*, 2399–2411.
- (27) Valverde, I. E.; Lecaille, F.; Lalmanach, G.; Aucagne, V.; Delmas, A. F. Synthesis of a biologically active triazole-containing analogue of cystatin A through successive peptidomimetic alkyne-azide ligations. *Angew. Chem., Int. Ed.* **2012**, *51*, 718–722.
- (28) Valverde, I. E.; Bauman, A.; Kluba, C. A.; Vomstein, S.; Walter, M. A.; Mindt, T. L. 1,2,3-triazoles as amide bond mimics: triazole scan yields protease-resistant peptidomimetics for tumor targeting. *Angew. Chem., Int. Ed.* **2013**, *52*, 8957–8960.
- (29) Wu, F. Z.; Mayer, J. P.; Gelfanov, V. M.; Liu, F.; DiMarchi, R. D. Synthesis of four-disulfide insulin analogs via sequential disulfide bond formation. *J. Org. Chem.* **2017**, *82*, 3506–3512.
- (30) Karas, J. A.; Patil, N. A.; Tailhades, J.; Sani, M. A.; Scanlon, D. B.; Forbes, B. E.; Gardiner, J.; Separovic, F.; Wade, J. D.; Hossain, M. A. Total chemical synthesis of an intra-A-chain cystathionine human insulin analogue with enhanced thermal stability. *Angew. Chem., Int. Ed.* **2016**, *55*, 14743–14747.
- (31) Liu, F.; Zaykov, A. N.; Levy, J. J.; DiMarchi, R. D.; Mayer, J. P. Chemical synthesis of peptides within the insulin superfamily. *J. Pept. Sci.* **2016**, *22*, 260–270.
- (32) Hossain, M. A.; Rosengren, K. J.; Zhang, S. D.; Bathgate, R. A. D.; Tregear, G. W.; van Lierop, B. J.; Robinson, A. J.; Wade, J. D. Solid phase synthesis and structural analysis of novel A-chain dicarba analogs of human relaxin-3 (INSL7) that exhibit full biological activity. *Org. Biomol. Chem.* **2009**, *7*, 1547–1553.
- (33) Zhang, S. D.; Hughes, R. A.; Bathgate, R. A. D.; Shabanpoor, F.; Hossain, M. A.; Lin, F.; van Lierop, B.; Robinson, A. J.; Wade, J. D. Role of the intra-A-chain disulfide bond of insulin-like peptide 3 in binding and activation of its receptor, RXFP2. *Peptides* **2010**, *31*, 1730–1736.
- (34) Shabanpoor, F.; Hossain, M. A.; Ryan, P. J.; Belgi, A.; Layfield, S.; Kocan, M.; Zhang, S. D.; Samuel, C. S.; Gundlach, A. L.; Bathgate, R. A. D.; Separovic, F.; Wade, J. D. Minimization of human relaxin-3 leading to high-affinity analogues with increased selectivity for relaxin-family peptide 3 receptor (RXFP3) over RXFP1. *J. Med. Chem.* **2012**, *55*, 1671–1681.
- (35) Hossain, M. A.; Haugaard-Kedstrom, L. M.; Rosengren, K. J.; Bathgate, R. A. D.; Wade, J. D. Chemically synthesized dicarba H2 relaxin analogues retain strong RXFP1 receptor activity but show an unexpected loss of in vitro serum stability. *Org. Biomol. Chem.* **2015**, *13*, 10895–10903.
- (36) Lin, F.; Hossain, M. A.; Post, S.; Karashchuk, G.; Tatar, M.; De Meyts, P.; Wade, J. D. Total solid-phase synthesis of biologically active drosophila insulin-like peptide 2 (DILP2). *Aust. J. Chem.* **2017**, *70*, 208–212.
- (37) Menting, J. G.; Gajewiak, J.; MacRaid, C. A.; Chou, D. H. C.; Disotuar, M. M.; Smith, N. A.; Miller, C.; Ercegyi, J.; Rivier, J. E.; Olivera, B. M.; Forbes, B. E.; Smith, B. J.; Norton, R. S.; Safavi-Hemami, H.; Lawrence, M. C. A minimized human insulin-receptor-binding motif revealed in a *Conus geographus* venom insulin. *Nat. Struct. Mol. Biol.* **2016**, *23*, 916–920.
- (38) Du, Y. C.; Zhang, Y. S.; Lu, Z. X.; Tsou, C. L. Resynthesis of insulin from its glycyl and phenylalanyl chains. *Sci. Sin.* **1961**, *10*, 84–104.
- (39) Ruegg, U. T.; Gattner, H. G. Reduction of S-sulfo groups by tributylphosphine - improved method for recombination of insulin chains. *Hoppe-Seyler's Z. Physiol. Chem.* **1975**, *356*, 1527–1533.
- (40) Chance, R. E.; Hoffmann, J. A.; Kroeff, E. P.; Johnson, M. G.; Schirmer, E. W.; Bromer, W. W. The Production of Human Insulin Using Recombinant DNA Technology and a New Chain Combination Procedure. In *Proceedings of the 7th American Peptide Symposium*;

- Rich, D. H., Gross, E., Eds.; Pierce Chemical Company: Rockford, IL, 1981; pp 721–728.
- (41) Wang, S. H.; Hu, S. Q.; Burke, G. T.; Katsouyannis, P. G. Insulin analogs with modifications in the beta-turn of the B-chain. *J. Protein Chem.* **1991**, *10*, 313–324.
- (42) Krizkova, K.; Veverka, V.; Maletinska, L.; Hexnerova, R.; Brzozowski, A. M.; Jiracek, J.; Zakova, L. Structural and functional study of the GlnB22-insulin mutant responsible for maturity-onset diabetes of the young. *PLoS One* **2014**, *9*, e112883.
- (43) Kosinova, L.; Veverka, V.; Novotna, P.; Collinsova, M.; Urbanova, M.; Moody, N. R.; Turkenburg, J. P.; Jiracek, J.; Brzozowski, A. M.; Zakova, L. Insight into the structural and biological relevance of the T/R transition of the N-terminus of the B-chain in human insulin. *Biochemistry* **2014**, *53*, 3392–3402.
- (44) Bayne, M. L.; Applebaum, J.; Underwood, D.; Chicchi, G. G.; Green, B. G.; Hayes, N. S.; Cascieri, M. A. The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor. *J. Biol. Chem.* **1989**, *264*, 11004–11008.
- (45) Gill, R.; Wallach, B.; Verma, C.; Urso, B.; DeWolf, E.; Grotzinger, J.; MurrayRust, J.; Pitts, J.; Wollmer, A.; DeMeyts, P.; Wood, S. Engineering the C-region of human insulin-like growth factor-1: implications for receptor binding. *Protein Eng., Des. Sel.* **1996**, *9*, 1011–1019.
- (46) Denley, A.; Bonython, E. R.; Booker, G. W.; Cosgrove, L. J.; Forbes, B. E.; Ward, C. W.; Wallace, J. C. Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. *Mol. Endocrinol.* **2004**, *18*, 2502–2512.
- (47) Zhang, W.; Gustafson, T. A.; Rutter, W. J.; Johnson, J. D. Positively charged side chains in the insulin-like growth factor-1 C- and D-regions determine receptor binding specificity. *Oncol. Rep.* **1994**, *269*, 10609–10613.
- (48) Jansson, M.; Andersson, G.; Uhlen, M.; Nilsson, B.; Kordel, J. The insulin-like growth factor (IGF) binding protein 1 binding epitope on IGF-I probed by heteronuclear NMR spectroscopy and mutational analysis. *J. Biol. Chem.* **1998**, *273*, 24701–24707.
- (49) Saegusa, J.; Yamaji, S.; Ieguchi, K.; Wu, C. Y.; Lam, K. S.; Liu, F. T.; Takada, Y. K.; Takada, Y. The direct binding of insulin-like growth factor-1 (IGF-1) to integrin $\alpha v \beta 3$ is involved in IGF-1 signaling. *J. Biol. Chem.* **2009**, *284*, 24106–24114.
- (50) Fujita, M.; Ieguchi, K.; Cedano-Prieto, D. M.; Fong, A.; Wilkerson, C.; Chen, J. Q.; Wu, M.; Lo, S. H.; Cheung, A. T. W.; Wilson, M. D.; Cardiff, R. D.; Borowsky, A. D.; Takada, Y. K.; Takada, Y. An integrin binding-defective mutant of insulin-like growth factor-1 (R36E/R37E IGF1) acts as a dominant-negative antagonist of the IGF1 receptor (IGF1R) and suppresses tumorigenesis but still binds to IGF1R. *J. Biol. Chem.* **2013**, *288*, 19593–19603.
- (51) Brzozowski, A. M.; Dodson, E. J.; Dodson, G. G.; Murshudov, G. N.; Verma, C.; Turkenburg, J. P.; de Bree, F. M.; Dauter, Z. Structural origins of the functional divergence of human insulin-like growth factor-I and insulin. *Biochemistry* **2002**, *41*, 9389–9397.
- (52) Vajdos, F. F.; Ultsch, M.; Schaffer, M. L.; Deshayes, K. D.; Liu, J.; Skelton, N. J.; de Vos, A. M. Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions. *Biochemistry* **2001**, *40*, 11022–11029.
- (53) Sitar, T.; Popowicz, G. M.; Siwanowicz, I.; Huber, R.; Holak, T. A. Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13028–13033.
- (54) Jiracek, J.; Zakova, L. Structural perspectives of insulin receptor isoform-selective insulin analogs. *Front. Endocrinol.* **2017**, *8*, 167.
- (55) Krizkova, K.; Chrudinova, M.; Povalova, A.; Selicharova, I.; Collinsova, M.; Vanek, V.; Brzozowski, A. M.; Jiracek, J.; Zakova, L. Insulin-insulin-like growth factors hybrids as molecular probes of hormone:receptor binding specificity. *Biochemistry* **2016**, *55*, 2903–2913.
- (56) Isaad, A. L.; Papini, A. M.; Chorev, M.; Rovero, P. Side chain-to-side chain cyclization by click reaction. *J. Pept. Sci.* **2009**, *15*, 451–454.
- (57) Vikova, J.; Collinsova, M.; Kletvikova, E.; Budesinsky, M.; Kaplan, V.; Zakova, L.; Veverka, V.; Hexnerova, R.; Avino, R. J. T.; Strakova, J.; Selicharova, I.; Vanek, V.; Wright, D. W.; Watson, C. J.; Turkenburg, J. P.; Brzozowski, A. M.; Jiracek, J. Rational steering of insulin binding specificity by intra-chain chemical crosslinking. *Sci. Rep.* **2016**, *6*, 19431.
- (58) Vanek, V.; Picha, J.; Fabre, B.; Budesinsky, M.; Lepsik, M.; Jiracek, J. The development of a versatile trifunctional scaffold for biological applications. *Eur. J. Org. Chem.* **2015**, *2015*, 3689–3701.
- (59) Fabre, B.; Picha, J.; Vanek, V.; Budesinsky, M.; Jiracek, J. A CuAAC-hydrazone-CuAAC trifunctional scaffold for the solid-phase synthesis of trimodal compounds: Possibilities and limitations. *Molecules* **2015**, *20*, 19310–19329.
- (60) Fabre, B.; Picha, J.; Vanek, V.; Selicharova, I.; Chrudinova, M.; Collinsova, M.; Zakova, L.; Budeginsky, M.; Jiracek, J. Synthesis and evaluation of a library of trifunctional scaffold-derived compounds as modulators of the insulin receptor. *ACS Comb. Sci.* **2016**, *18*, 710–722.
- (61) Isidro-Llobet, A.; Alvarez, M.; Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **2009**, *109*, 2455–2504.
- (62) Behrendt, R.; White, P.; Offer, J. Advances in Fmoc solid-phase peptide synthesis. *J. Pept. Sci.* **2016**, *22*, 4–27.
- (63) Subiros-Funosas, R.; El-Faham, A.; Albericio, F. Aspartimide formation in peptide chemistry: occurrence, prevention strategies and the role of N-hydroxylamines. *Tetrahedron* **2011**, *67*, 8595–8606.
- (64) Behrendt, R.; Willibald, J.; White, P. New Aspartic Acid Derivatives Useful for Synthesizing Peptide Which is Useful for Treatment of Crohn's Disease and Osteoporosis and as Adjuvant Therapy in Chemotherapy. European patent EP2886531, 2015.
- (65) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem., Int. Ed.* **2009**, *48*, 9879–9883.
- (66) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.* **2004**, *6*, 2853–2855.
- (67) Horne, W. S.; Yadav, M. K.; Stout, C. D.; Ghadiri, M. R. Heterocyclic peptide backbone modifications in an alpha-helical coiled coil. *J. Am. Chem. Soc.* **2004**, *126*, 15366–15367.
- (68) Shooter, G. K.; Magee, B.; Soos, M. A.; Francis, G. L.; Siddle, K.; Wallace, J. C. Insulin-like growth factor (IGF)-I A- and B-domain analogues with altered type 1 IGF and insulin receptor binding specificities. *J. Mol. Endocrinol.* **1996**, *17*, 237–246.
- (69) Menting, J. G.; Whittaker, J.; Margetts, M. B.; Whittaker, L. J.; Kong, G. K. W.; Smith, B. J.; Watson, C. J.; Zakova, L.; Kletvikova, E.; Jiracek, J.; Chan, S. J.; Steiner, D. F.; Dodson, G. G.; Brzozowski, A. M.; Weiss, M. A.; Ward, C. W.; Lawrence, M. C. How insulin engages its primary binding site on the insulin receptor. *Nature* **2013**, *493*, 241–245.
- (70) Menting, J. G.; Lawrence, C. F.; Kong, G. K. W.; Margetts, M. B.; Ward, C. W.; Lawrence, M. C. Structural congruency of ligand binding to the insulin and insulin/type 1 insulin-like growth factor hybrid receptors. *Structure* **2015**, *23*, 1271–1282.
- (71) De Meyts, P. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling. *Diabetologia* **1994**, *37*, S135–S148.
- (72) Picha, J.; Buděšínský, M.; Machackova, K.; Collinsova, M.; Jiracek, J. Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides. *J. Pept. Sci.* **2017**, *23*, 202–214.
- (73) Sell, C.; Dumenil, G.; Deveaud, C.; Miura, M.; Coppola, D.; Deangelis, T.; Rubin, R.; Efstratiadis, A.; Baserga, R. Effect of a null mutation of the insulin-like growth-factor-I receptor gene on growth and transformation of mouse embryo fibroblasts. *Mol. Cell. Biol.* **1994**, *14*, 3604–3612.
- (74) Slaaby, R.; Andersen, A. S.; Brandt, J. IGF-I binding to the IGF-I receptor is affected by contaminants in commercial BSA: The contaminants are proteins with IGF-I binding properties. *Growth Horm. IGF Res.* **2008**, *18*, 267–274.
- (75) Aguilar, H. N.; Zielnik, B.; Tracey, C. N.; Mitchell, B. F. Quantification of rapid myosin regulatory light chain phosphorylation

using high-throughput in-cell western assays: comparison to western immunoblots. *PLoS One* **2010**, *5*, e9965.

Supporting Information
for
**INSULIN-LIKE GROWTH FACTOR 1 ANALOGS CLICKED IN THE C DOMAIN:
CHEMICAL SYNTHESIS AND BIOLOGICAL ACTIVITIES**

Kateřina Macháčková¹, Michaela Collinsová¹, Martina Chrudinová¹, Irena Selicharová¹,
Jan Pícha¹, Miloš Buděšínský¹, Václav Vaněk¹, Lenka Žáková¹, Andrzej M. Brzozowski²
and Jiří Jiráček^{1,*}

Table of Contents	Page 1
Figure S1. Schematic overview of the synthetic pathways leading to analog 1	Page 3
Figure S2. Schematic overview of the synthetic pathways leading to analog 2	Page 4
Figure S3. Schematic overview of the synthetic pathways leading to analog 3	Page 5
Figure S4. Schematic overview of the synthetic pathways leading to analog 4	Page 6
Figure S5. Binding curves of analogs for IGF-1R	Page 7
Figure S6. Binding curves of analogs for IR-A	Page 8
RP-HPLC purification and analyses of compounds	Page 9
Figure S7. RP-HPLC analysis of the purified peptide 1a	Page 9
Figure S8. RP-HPLC analysis of the purified peptide 1b	Page 10
Figure S9. RP-HPLC purification of analog 1	Page 10
Figure S10. RP-HPLC analysis of the purified analog 1	Page 11
Figure S11. HRMS-ESI spectrum of purified analog 1	Page 12
Figure S12. RP-HPLC purification of the crude peptide 2a prepared using Fmoc-Asp(OtBu) derivative	Page 13
Figure S13. RP-HPLC analysis of the purified peptide 2a prepared using Fmoc-Asp(OEpe) derivative	Page 13
Figure S14. RP-HPLC analysis of the purified peptide 2b	Page 14
Figure S15. RP-HPLC analysis of the purified analog 2c	Page 14
Figure S16. RP-HPLC purification of analog 2	Page 15
Figure S17. RP-HPLC analysis of the purified analog 2	Page 16
Figure S18. HRMS-ESI spectrum of purified analog 2	Page 17

Figure S19. RP-HPLC analysis of the purified peptide 3a	Page 18
Figure S20. RP-HPLC analysis of the purified peptide 3b	Page 18
Figure S21. RP-HPLC analysis of the purified analog 3	Page 19
Figure S22. HR-ME ESI spectrum of the purified analog 3	Page 20
Figure S23. RP-HPLC analysis of the purified peptide 4a	Page 21
Figure S24. RP-HPLC analysis of the purified peptide 4b	Page 21
Figure S25. RP-HPLC analysis of the purified analog 4c	Page 22
Figure S26. RP-HPLC analysis of the purified analog 4	Page 22
Figure S27. HRMS-ESI spectrum of purified analog 4	Page 23
Characterization of an aspartimide derivative of peptide 2a	Page 24
Figure S28. RP-HPLC analysis of the purified peptide 2ax	Page 24
NMR characterization of peptide 2ax	Page 25
Table S1. Proton chemical shifts of peptide 2ax in d ₆ DMSO at 310 K	Page 26
Synthesis and characterization of a model heptapeptide and its aspartimide derivative	Page 27
Figure S29. RP-HPLC analysis of the crude peptide 7	Page 27
Figure S30. Optimized RP-HPLC analysis of the crude peptide 7	Page 28
Table S2. Proton NMR data of peptides 7 , 7xa , 7xb and 7xc in DMSO at 310 K	Page 29
Table S3. Carbon-13 chemical shifts of peptides 7 , 7xa , 7xb and 7xc in DMSO at 310 K	Page 30
Scheme S1. Deduced structures of compounds 7 , 7xa , 7xb and 7xc	Page 31

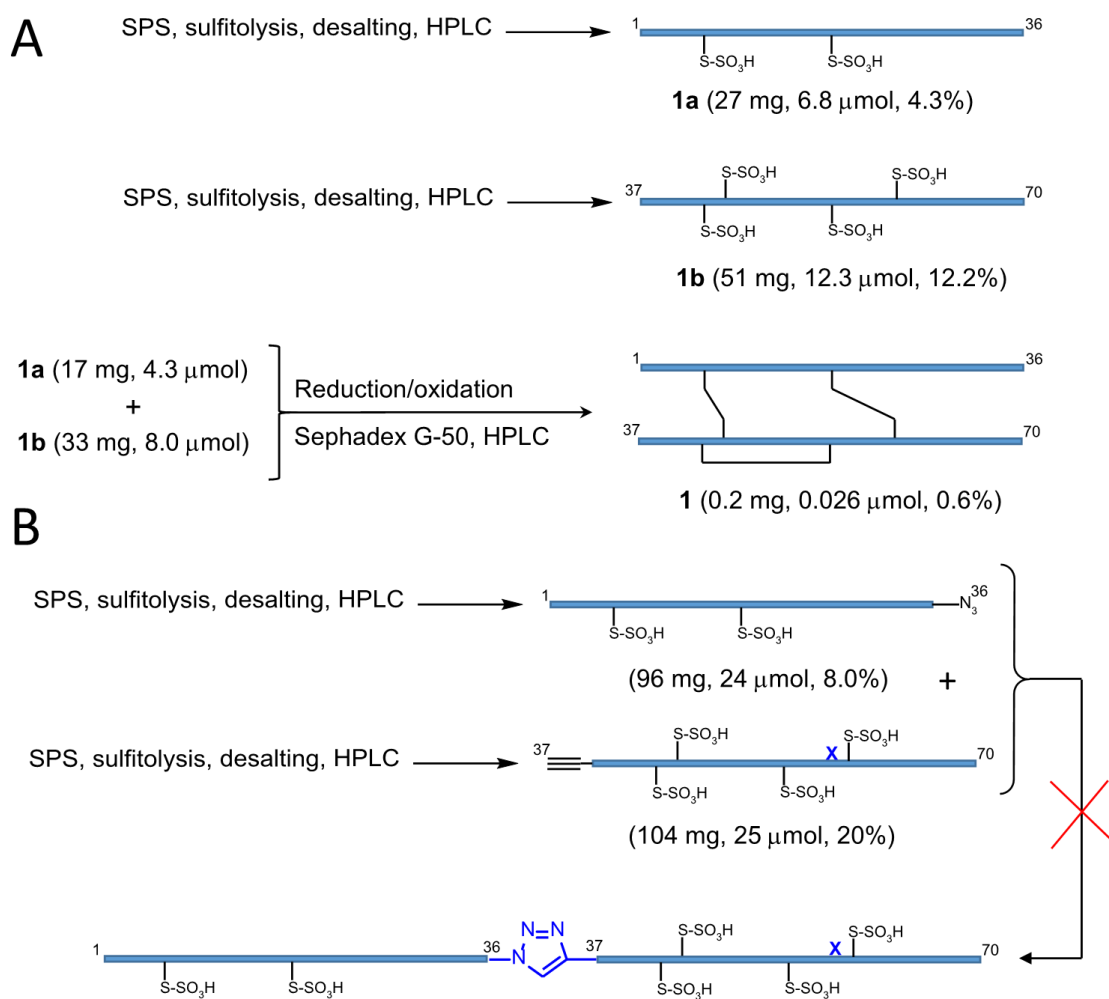


Figure S1. A. Schematic overview of the synthetic procedures for peptides **1a**, **1b** and analog **1**. **B** An overview of the unsuccessful synthesis of *S*-sulfonate protected IGF-1 chain connected by a triazole bridge between positions 36 and 37. Arginines 36 and 37 in *S*-sulfonate protected precursors were replaced by the C-terminal β-azido-alanine (position 36) and the N-terminal propargyl glycine (position 37). Substitution of Met59 for norleucine is shown as a blue X. SPS = solid-phase synthesis.

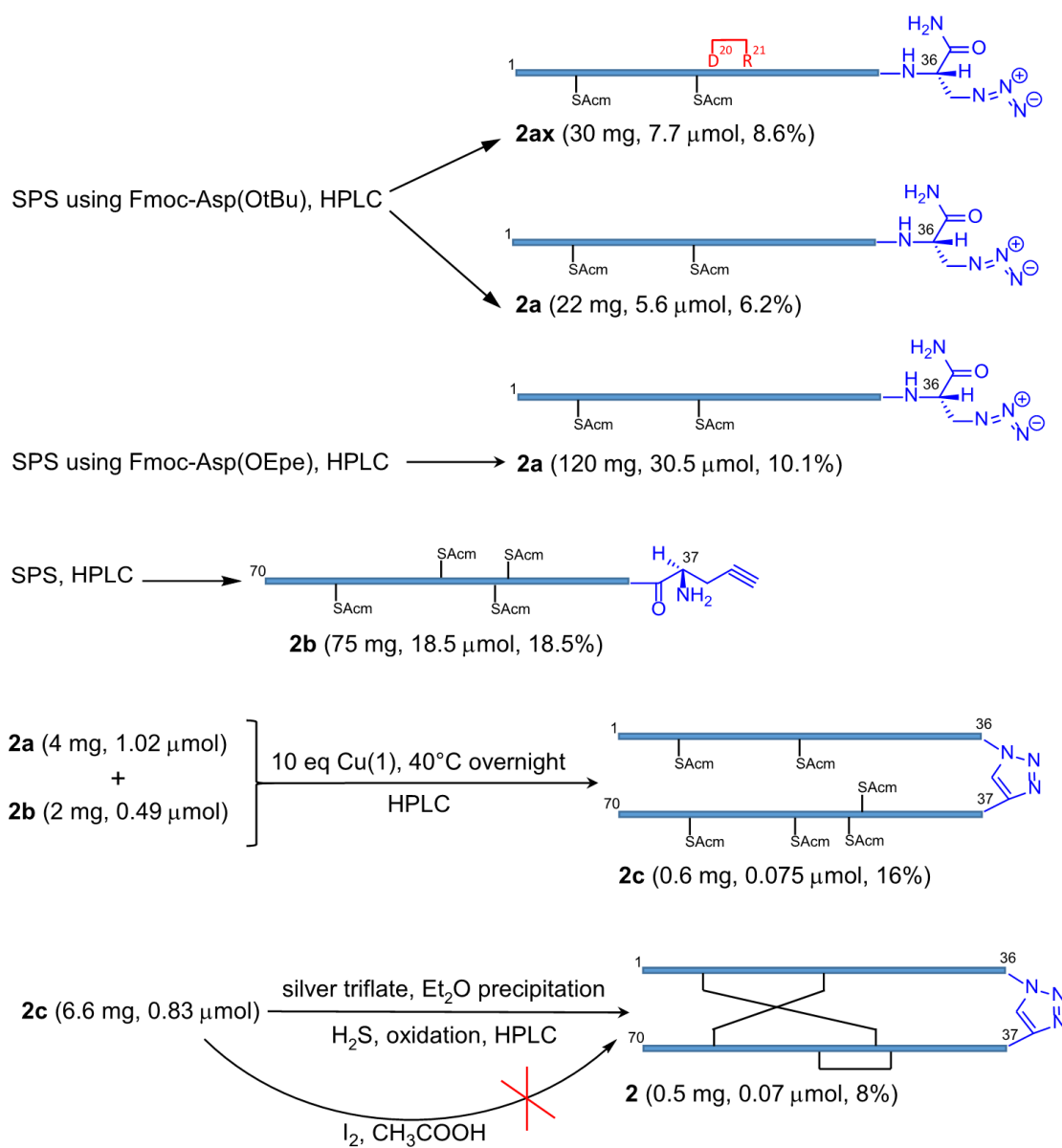


Figure S2. Schematic overview of the synthetic procedures for peptides **2a** (both synthetic variants, i.e. with Fmoc-Asp(OEpe) and Fmoc-Asp(OtBu), are shown), **2b**, **2c** and analog **2**. Undesired by-product **2ax** contains an aspartimide bond at Asp20-Arg21 (in red). Arginines 36 and 37 in S-Acm

protected precursors were replaced by the C-terminal β -azido-alanine (position 36) and the N-terminal propargyl glycine (position 37). Unsuccessful deprotection of AcM protecting groups from **2c** by iodine mentioned in the main text is shown as well. SPS = solid-phase synthesis.

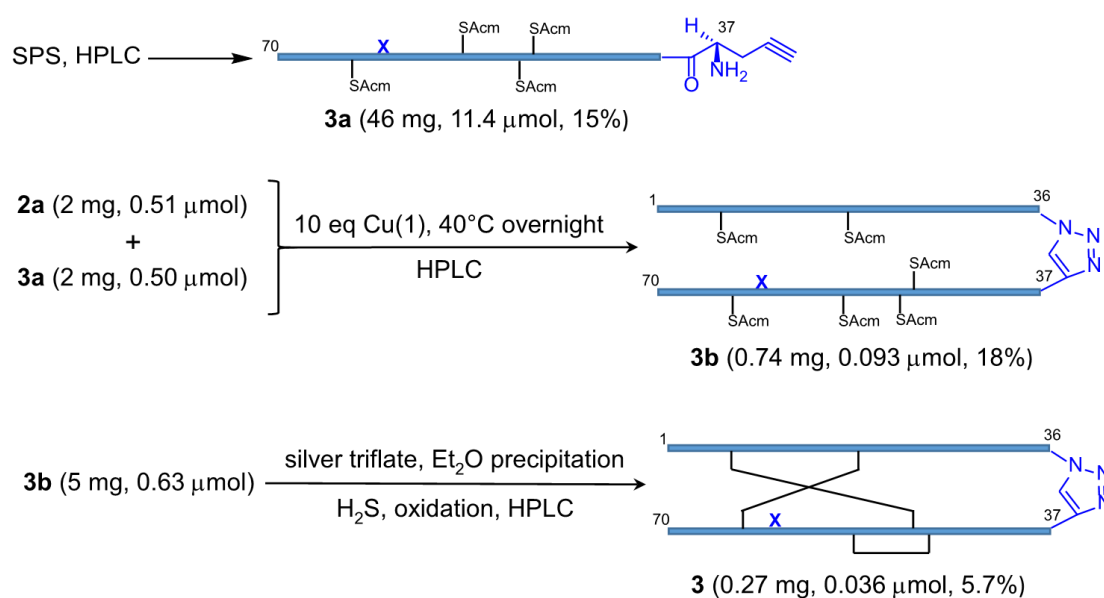


Figure S3. Schematic overview of the synthetic procedures for peptides **3a**, **3b** and analog **3**. Arginines 36 and 37 in *S*-AcM protected precursors were replaced by the C-terminal β -azido-alanine (position 36) and the N-terminal propargyl glycine (position 37). Substitution of Met59 for norleucine is shown as a blue X. SPS = solid-phase synthesis.

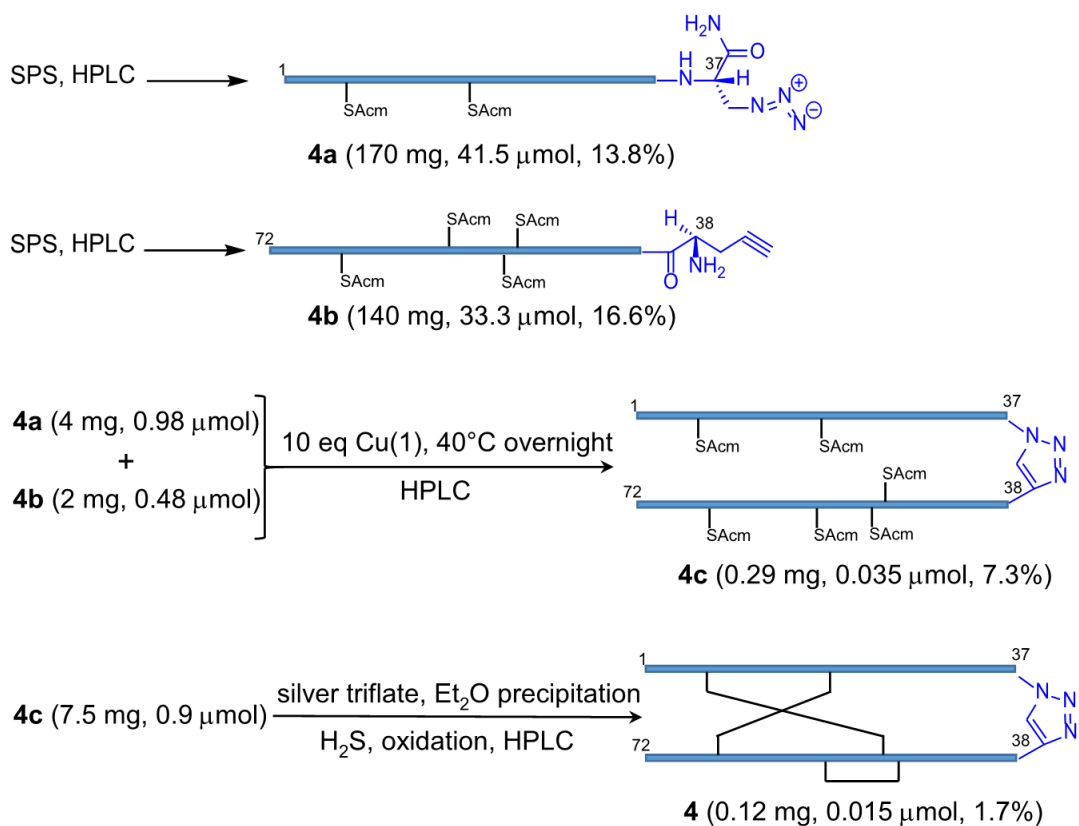


Figure S4. Schematic overview of the synthetic procedures for peptides **4a**, **4b**, **4c** and analog **4**. β -Azido-alanine was placed in the position 37 at the C-terminal of the precursor **4a**, and propargyl glycine was placed in the position 38 at the N-terminal of the precursor **4b**. SPS = solid-phase synthesis.

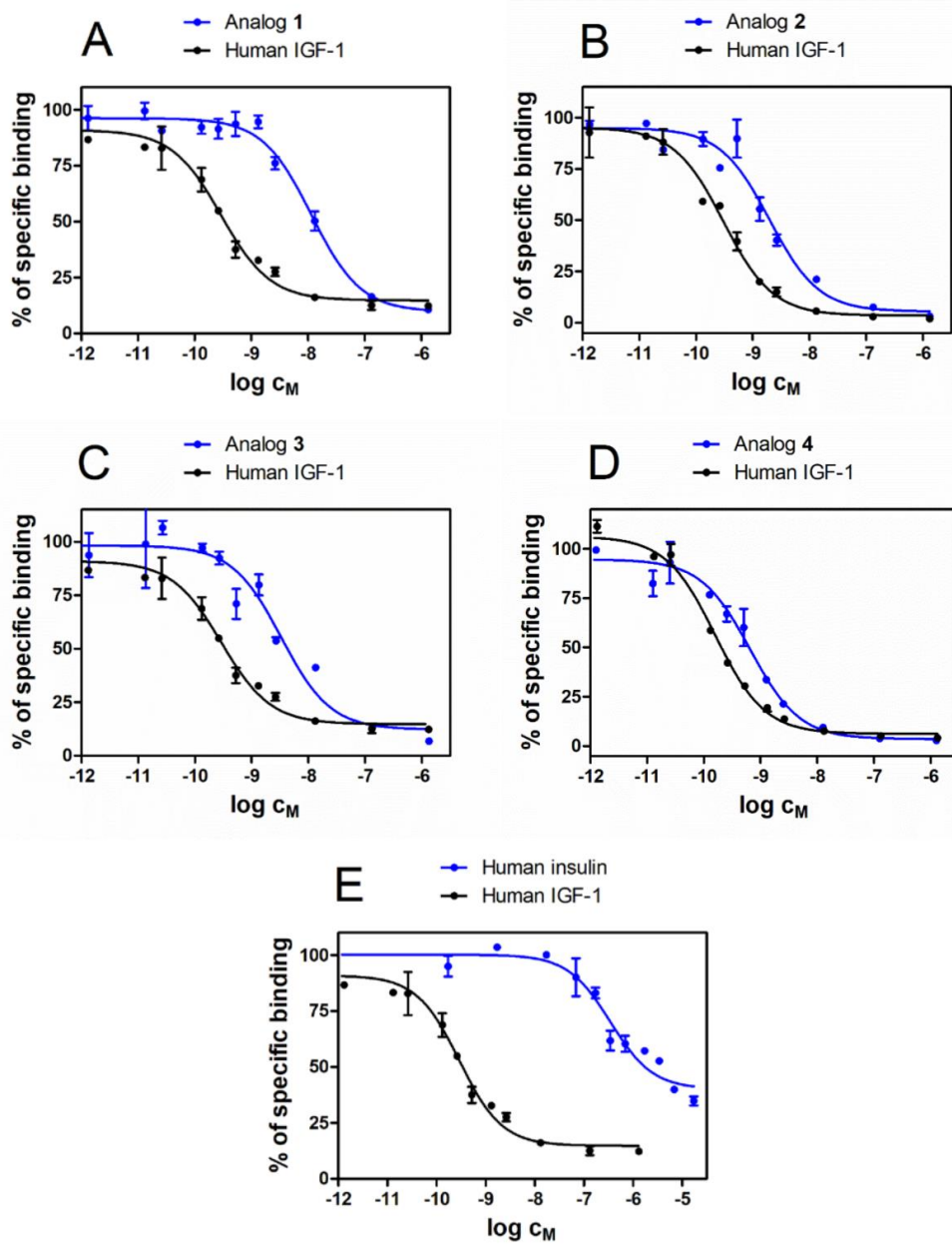


Figure S5. Inhibition of binding of human [¹²⁵I]moniodotyrosyl-IGF-1 to IGF-1R in membranes of mouse embryonic fibroblasts by human IGF-1, human insulin and IGF-1 analogs. Human IGF-1 is in black, analogs 1-4 or human insulin are in blue. A. Analog 1. B. Analog 2. C. Analog 3. D. Analog 4. E. Human insulin.

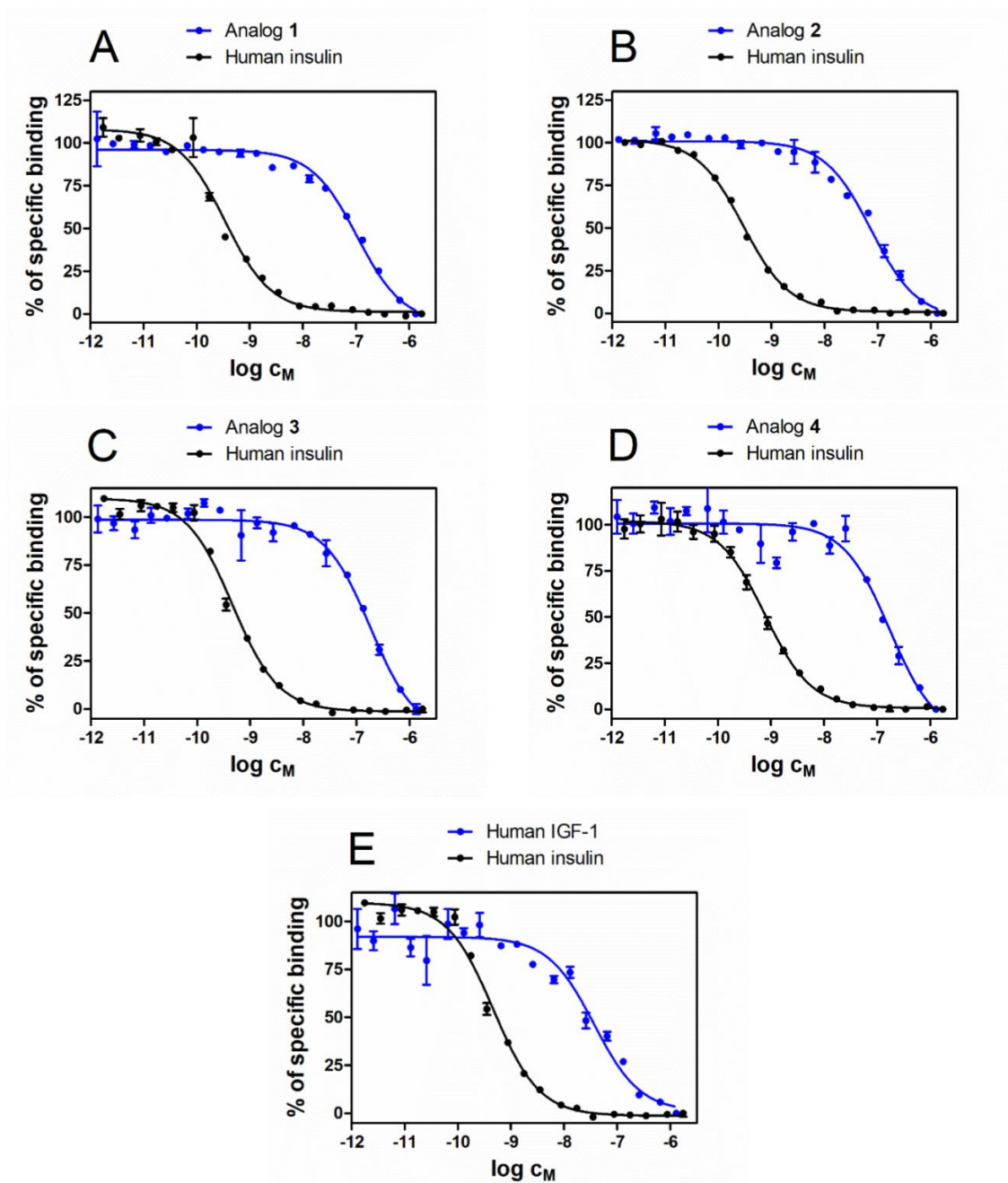


Figure S6. Inhibition of binding of human $[^{125}\text{I}]$ monoiodotyrosyl-insulin to IR-A in membranes of human IM-9 lymphocytes by human insulin, IGF-1 analogs and human IGF-1. Human IGF-1 is in black, analogs 1-4 or human insulin are in blue. A. Analog 1. B. Analog 2. C. Analog 3. D. Analog 4. E. Human IGF-1.

RP-HPLC purification and analyses of compounds

Precursor peptides were purified using reverse-phase high-performance liquid chromatography (RP-HPLC) on a preparative Vydac C4 column (214TP101522, 10-15 μm , 250 x 22 mm) at 10 ml/min.

IGF-1 analogs were purified using reverse-phase high-performance liquid chromatography (RP-HPLC) on a semipreparative Vydac C4 column (214TP510, 5 μm , 250 x 10 mm) at 4 ml/min.

The purity of all compounds was verified on an analytical Vydac C4 column (214TP54, 5 μm , 250 x 4.6 mm) or on an analytical Macherey-Nagel C18 column (EC280/4 Nucleosil, 120-5, 250 x 4.6 mm) at 1 ml/min.

We used different gradients of acetonitrile in water and 0.1 % TFA (v/v), see below.

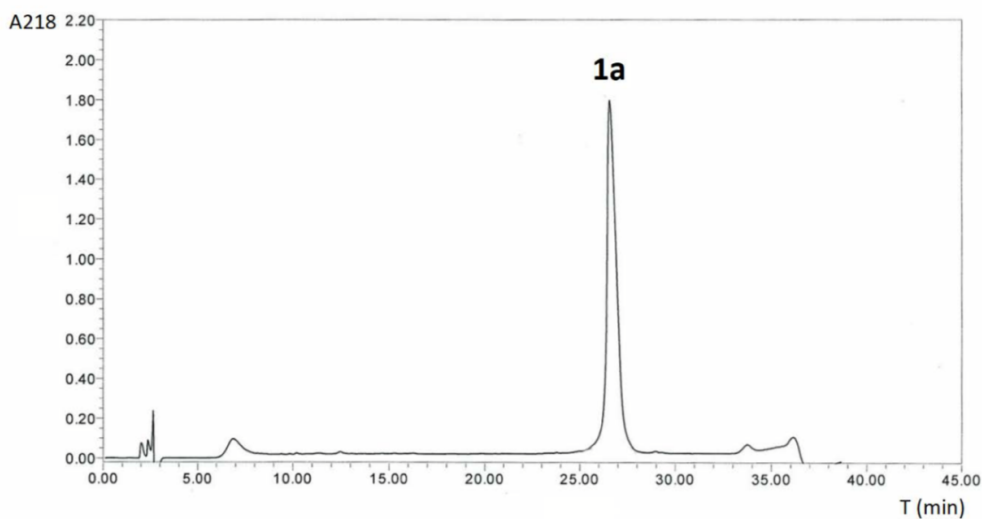


Figure S7. RP-HPLC analysis of the purified peptide **1a**. Gradient (min/% ACN): 0 /8 %, 1/28 %, 21/36 %, 34/44 %, 36/72 %, 37/72 %, 37.1/8 %. Macherey-Nagel C18 column (EC280/4 Nucleosil, 120-5, 250 x 4.6 mm).

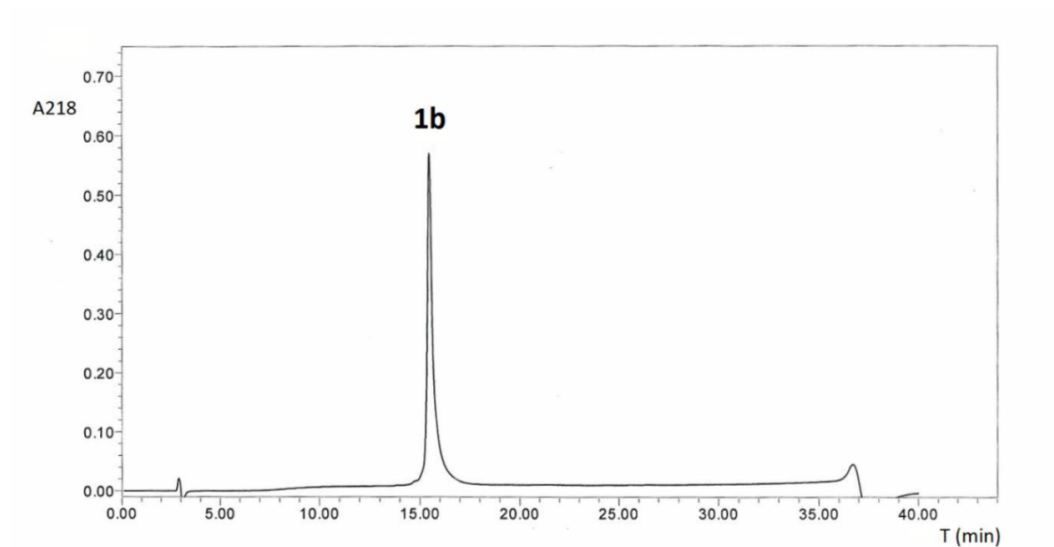


Figure S8. RP-HPLC analysis of the purified peptide **1b**. Gradient (min/% ACN): 0 /8 %, 30/80 %, 31/8 %. Macherey-Nagel C18 column (EC280/4 Nucleosil, 120-5, 250 x 4.6 mm).

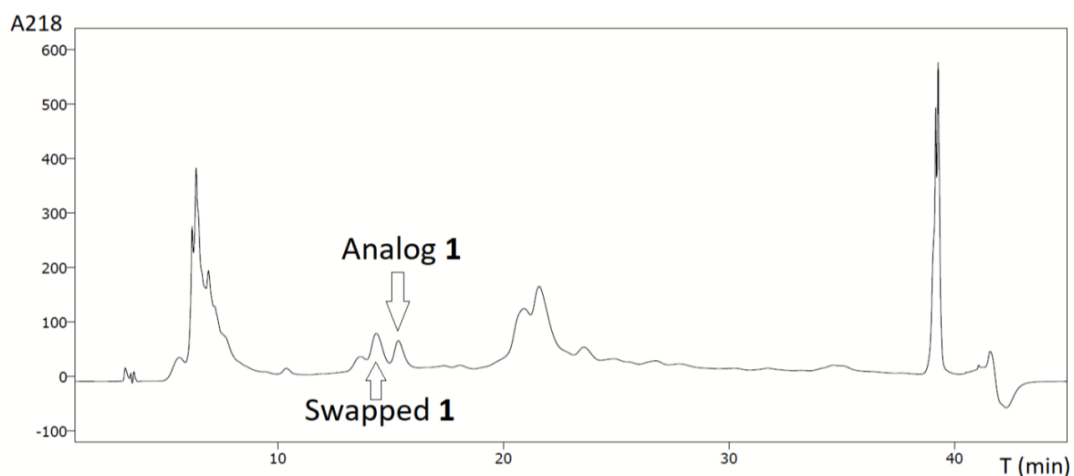


Figure S9. RP-HPLC purification of analog **1** after oxidative formation of disulfide bonds between precursors **1a** and **1b**. Gradient (min/% ACN): 0 /8 %, 1/27, 34/32, 36/80 %, 37/80, 37.1/8 %. Vydac C4 column (214TP510, 5 μ m, 250 x 10 mm). The peak with the analog **1** is marked by an arrow and the peak with the misfolded (swapped) product of **1** is marked as well.

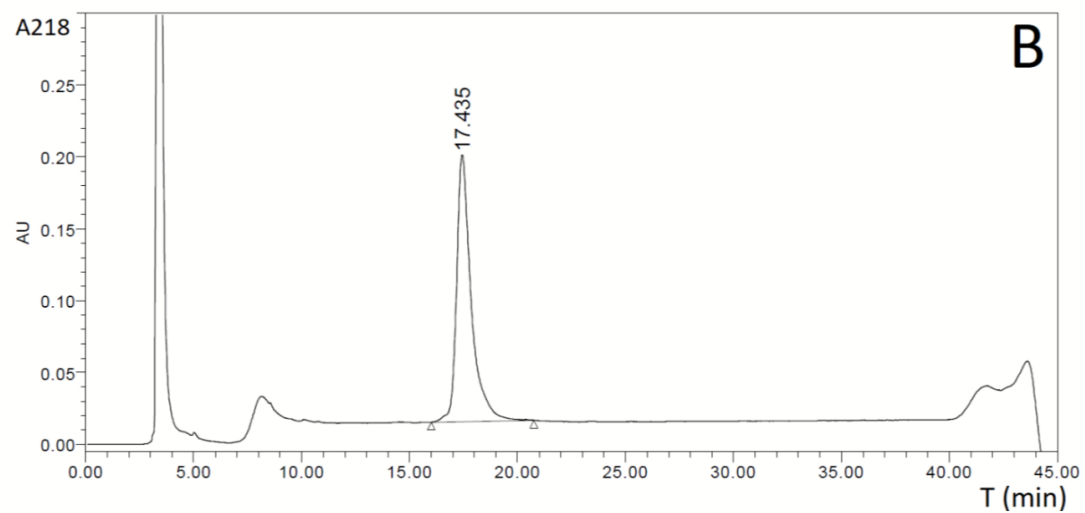
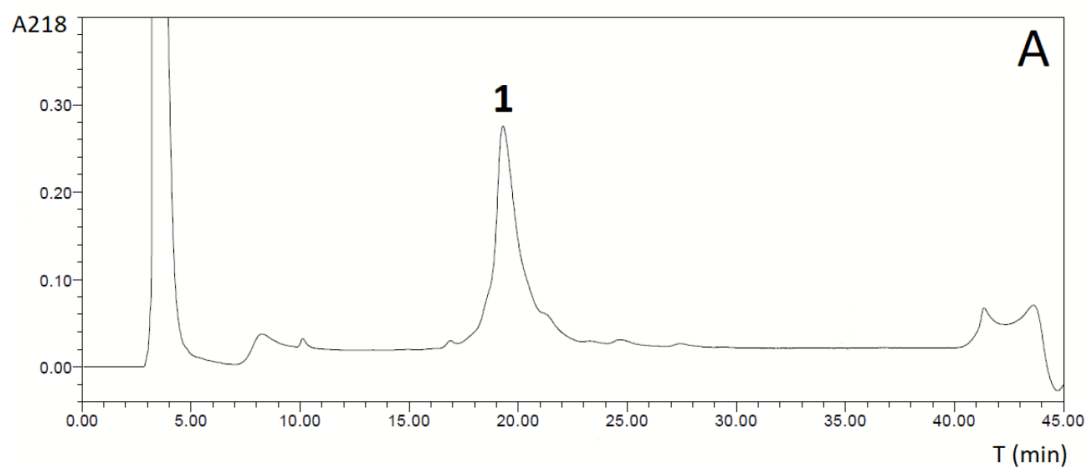


Figure S10. A. RP-HPLC analysis of the purified analog **1**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/8 %, 1/27 %, 34/64 %, 36/80 %, 37/80 %, 37.1/ 8%. **B.** RP-HPLC analysis of native human IGF-1 (Tercica) performed under the same conditions. The peak at 4 min. is caused by elution of acetic acid, in which the proteins were injected.

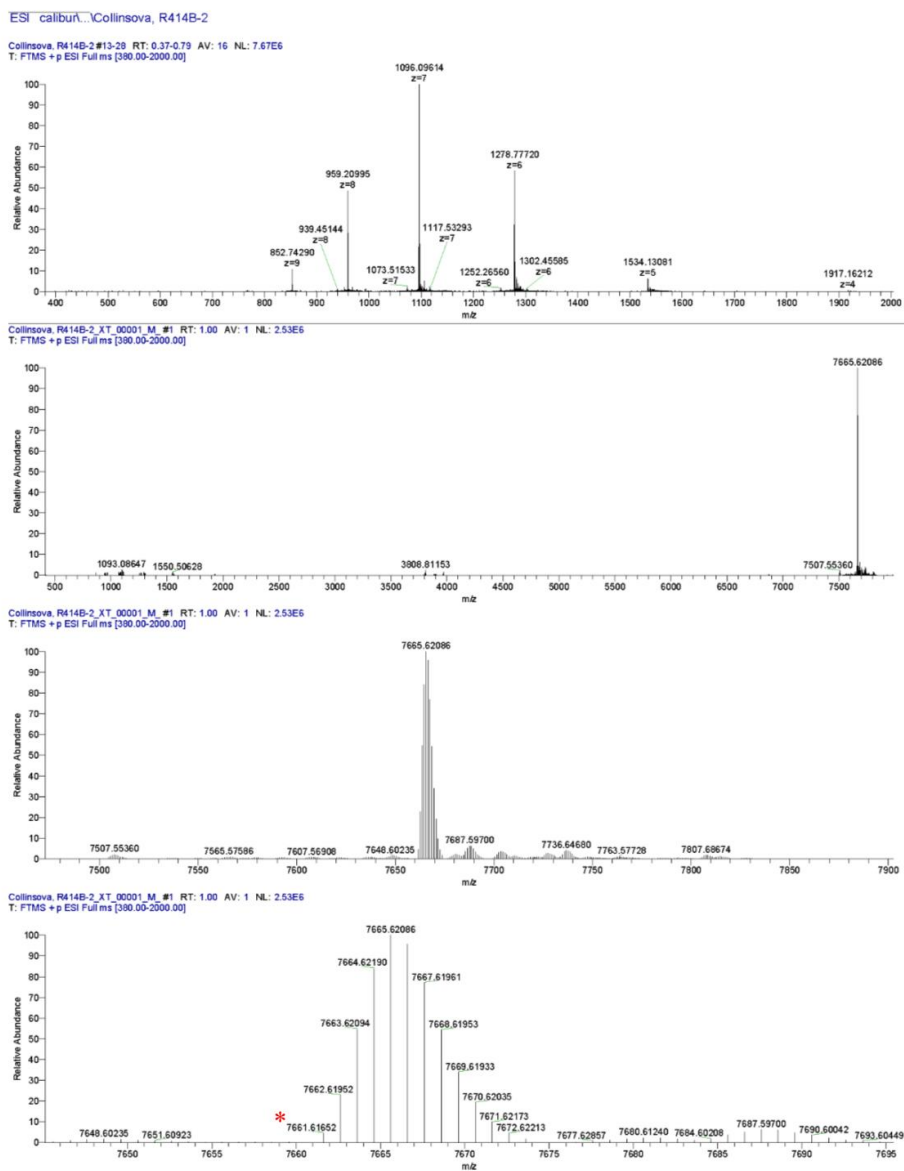


Figure S11. High-resolution ESI MS spectrum of purified analog **1**. Monoisotopic relative molecular weight (M_r) was calculated for $C_{331}H_{514}N_{94}O_{102}S_7$ and is 7661.597. The experimental monoisotopic M_r value was 7661.616 and the respective signal in the spectrum is marked by an asterisk.

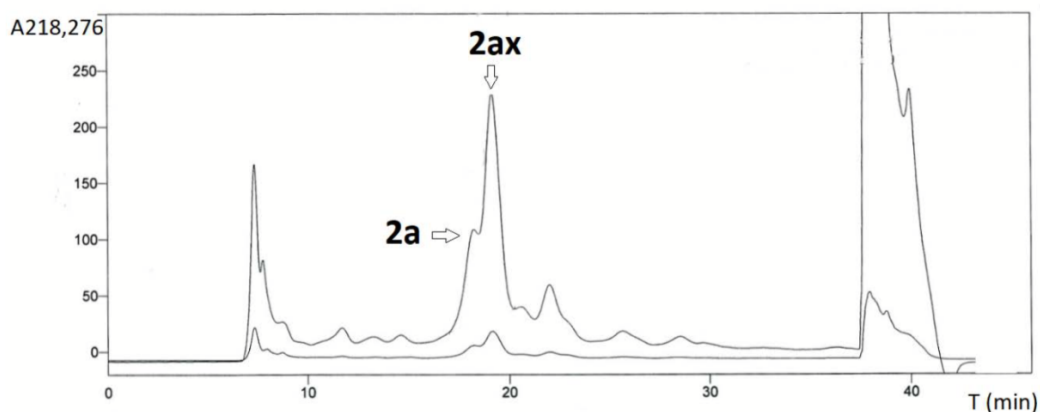


Figure S12. RP-HPLC purification of the crude peptide **2a** prepared using Fmoc-Asp(OtBu) derivative. The peptides in the peaks **2a** and **2ax** were identified as the target peptide and its aspartimide derivative (at Asp20-Arg21, see below), respectively. Vydac C4 column (214TP101522, 10-15 μm , 250 x 22 mm). Gradient (min/% ACN): 0/34 %, 30/39 %, 31/80 %, 32/80 %, 33/34 %.

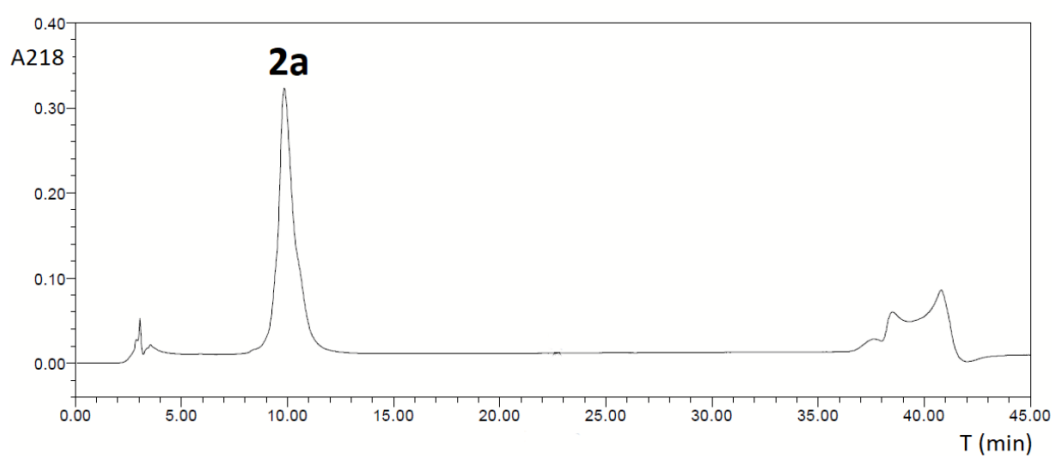


Figure S13. RP-HPLC analysis of the purified peptide **2a** prepared using Fmoc-Asp(OEpe) derivative. Vydac C4 column (214TP54, 5 μm , 250 x 4.6 mm). Gradient (min/% ACN): 0/33 %, 30/35 %, 31/80 %, 32/80 %, 33/33 %.

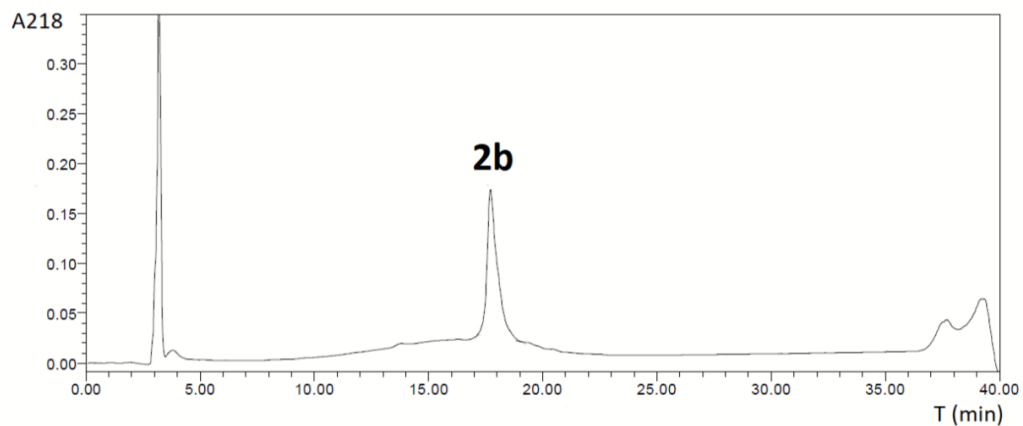


Figure S14. RP-HPLC analysis of the purified peptide **2b**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/22 %, 30/34 %, 31/80 %, 32/80 %, 33/22 %. The peak at 4 min. is caused by elution of acetic acid, in which the analog was injected.

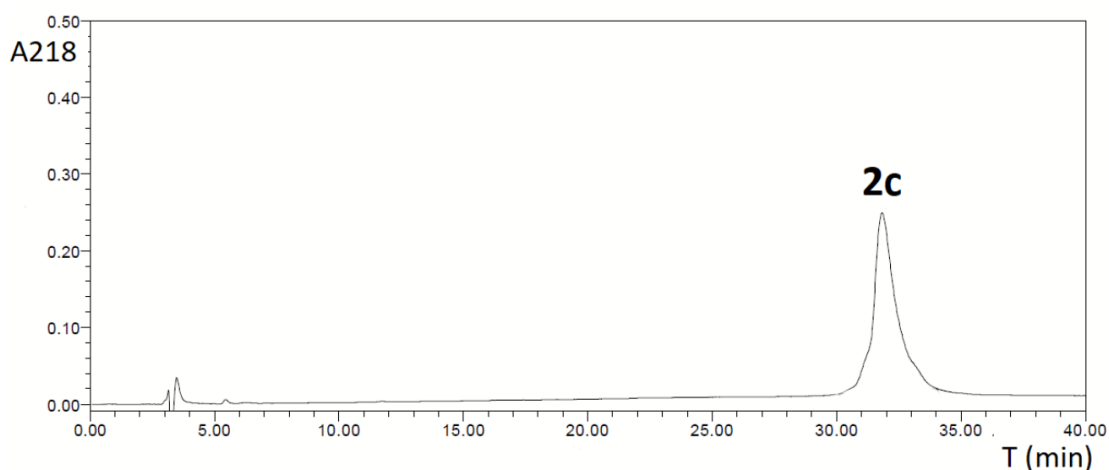


Figure S15. RP-HPLC analysis of the purified peptide **2c**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 20/32 %, 40/35 %, 41-42/80 %, 43/20 %. The peak at 4 min. is caused by elution of acetic acid, in which the analog was injected.

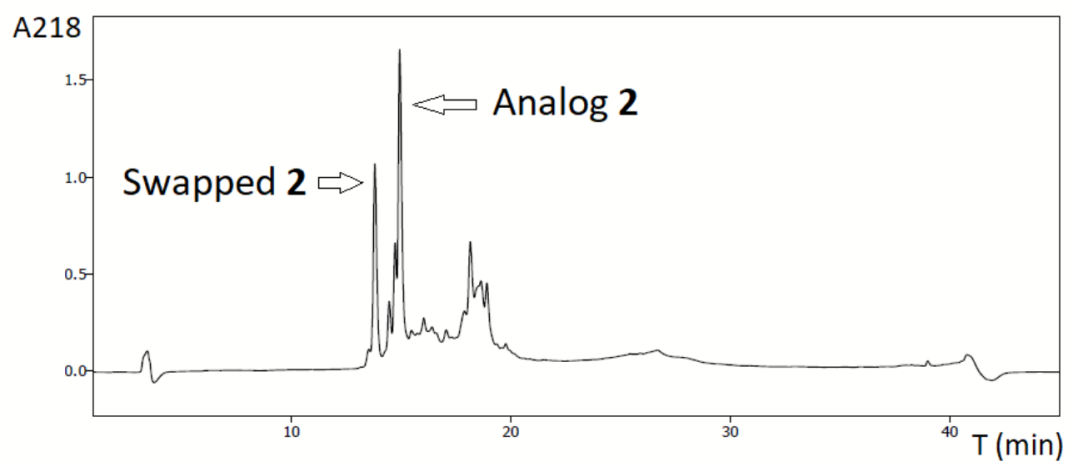


Figure S16. RP-HPLC purification of analog **2** after oxidative formation of disulfide bonds in the precursor **2c**. Gradient (min/% ACN): 0 /20 %, 33/44, 35-37/80, 37.1/20 %. Vydac C4 column (214TP510, 5 μ m, 250 x 10 mm). The peak with the analog **2** is marked by an arrow and the peak with the misfolded (swapped) product of **2** is marked as well. RP-HPLC reaction profiles for analogs **3** and **4** were similar (not shown).

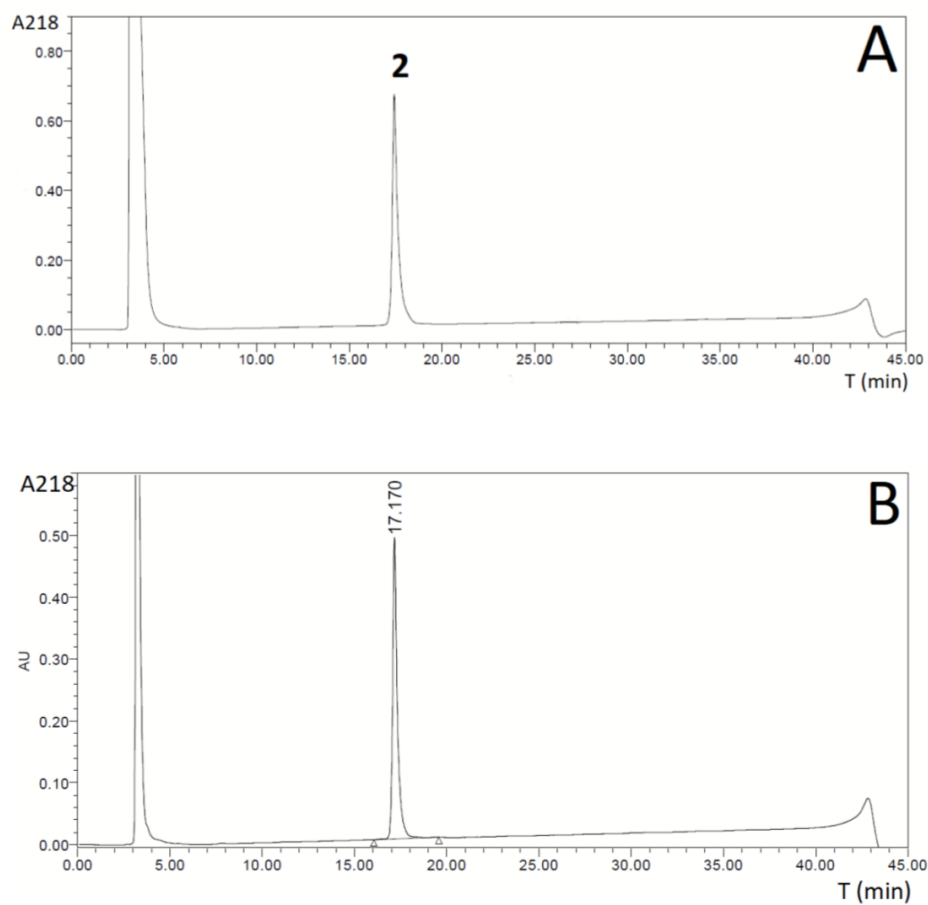


Figure S17. A. RP-HPLC analysis of the purified analog **2**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 33/54 %, 34/80 %, 35/80 %, 37/20 %. **B.** RP-HPLC analysis of native human IGF-1 (Tercica) performed under the same conditions. The peak at 4 min. is caused by elution of acetic acid, in which the proteins were injected.

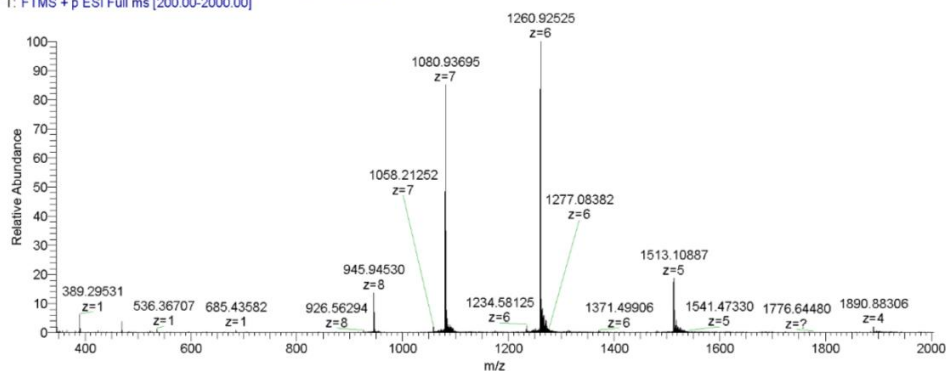
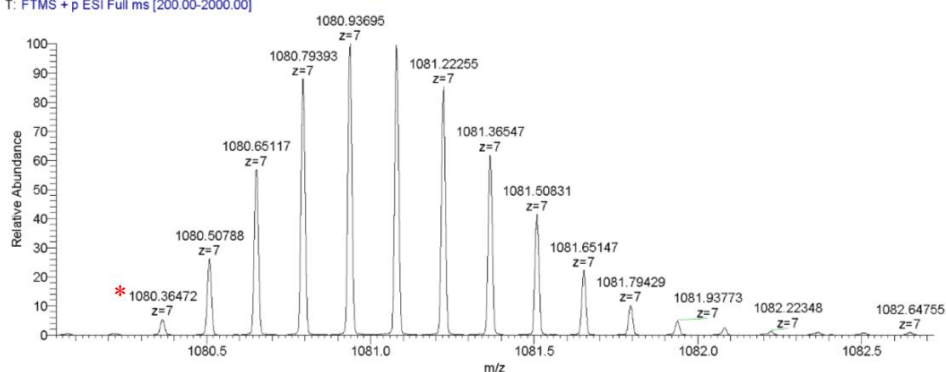
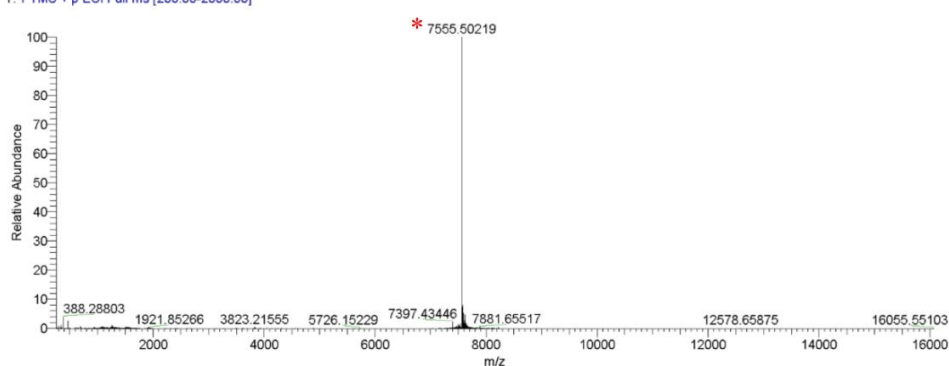
090916servisHR_5+ #59-70 RT: 1.64-1.96 AV: 12 NL: 2.88E6
T: FTMS + p ESI Full ms [200.00-2000.00]090916servisHR_5+ #59-70 RT: 1.64-1.96 AV: 12 NL: 2.45E6
T: FTMS + p ESI Full ms [200.00-2000.00]090916servisHR_5+ XT_00001_M #2 RT: 2.00 AV: 1 NL: 5.63E6
T: FTMS + p ESI Full ms [200.00-2000.00]

Figure S18. High-resolution ESI MS spectrum of purified analog **2**. Monoisotopic relative molecular weight (M_r) was calculated for $C_{327}H_{500}N_{92}O_{101}S_7$ and is 7555.486. The experimental monoisotopic M_r value was 7555.502 and the respective signal in the deconvoluted spectrum is marked by an asterisk.

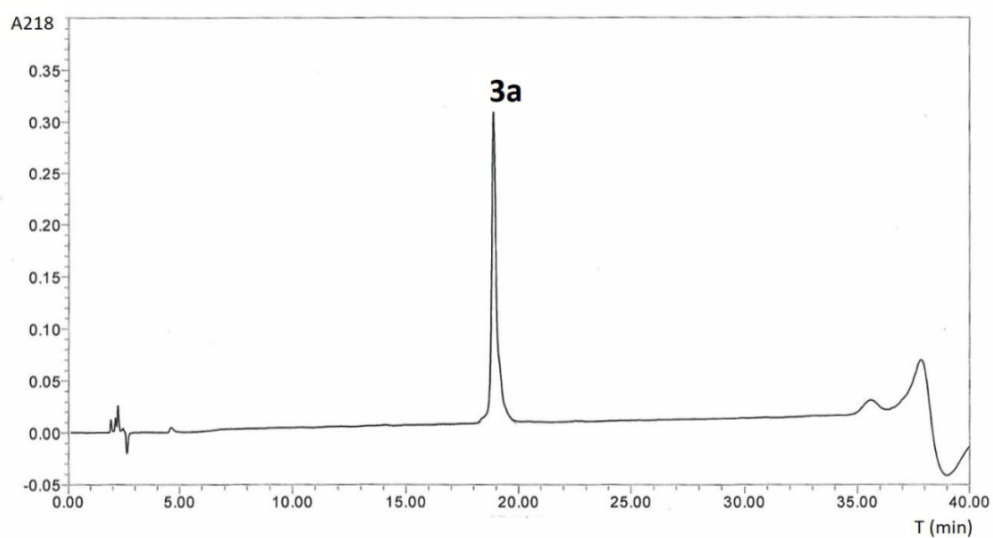


Figure S19. RP-HPLC analysis of the purified peptide **3a**. Vydac C4 column (214TP54, 5 μm , 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 30/44 %, 31/80 %, 32/80 %, 33/20 %.

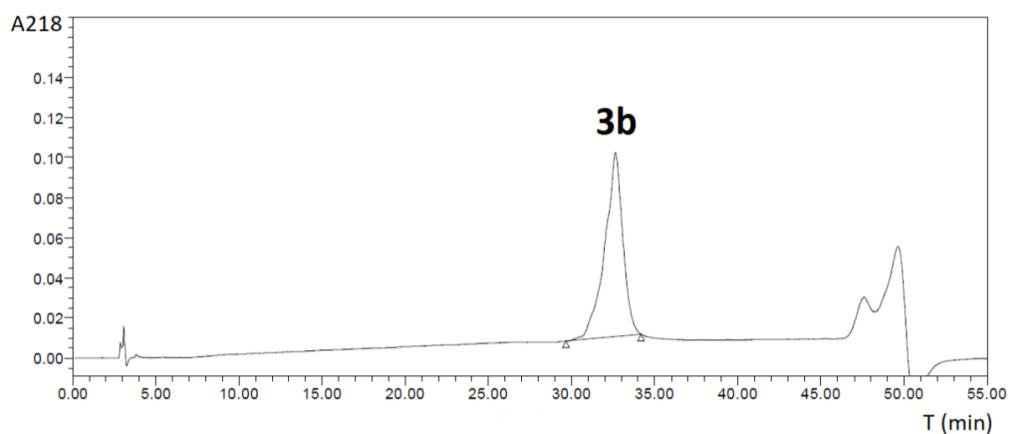


Figure S20. RP-HPLC analysis of the purified peptide **3b**. Vydac C4 column (214TP54, 5 μm , 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 20/32 %, 40/35 %, 41/80 %, 42/80 %, 43/20%.

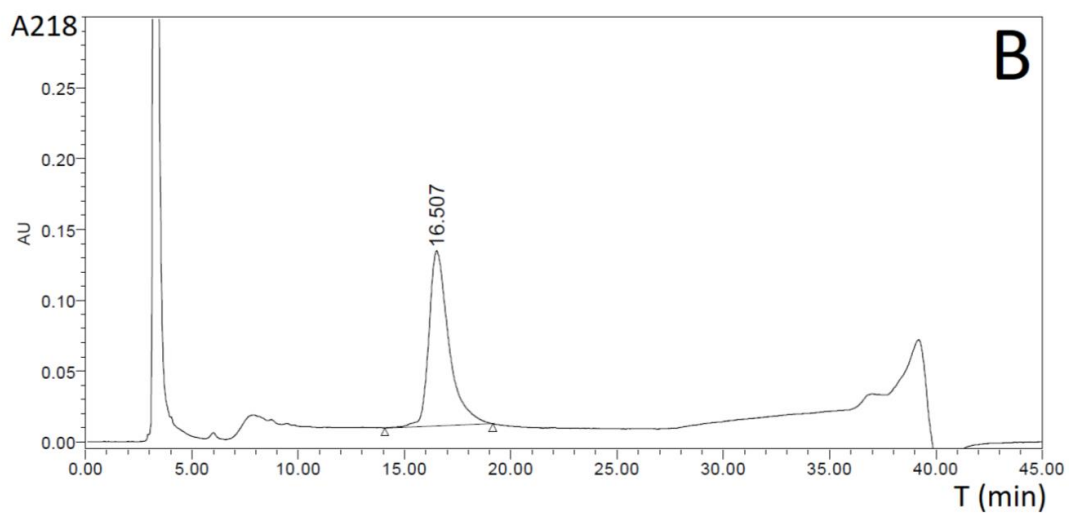
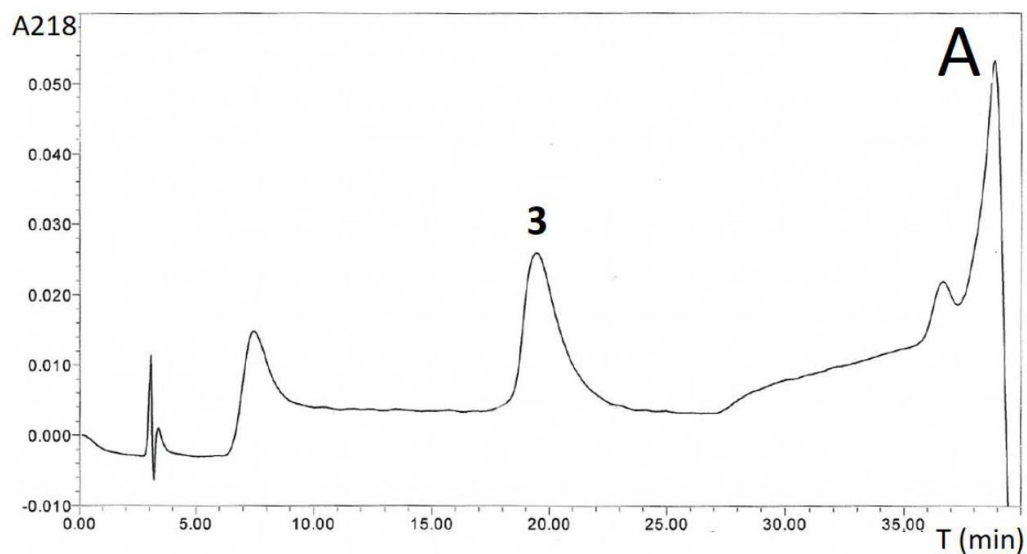


Figure S21. A. RP-HPLC analysis of the purified analog **3**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/16 %, 1/28 %, 21/28 %, 30/44 %, 31/80 %, 32/80%, 33/16%. **B.** RP-HPLC analysis of native human IGF-1 (Tercica) performed under the same conditions. The peak at 4 min. is caused by elution of acetic acid, in which the protein was injected.

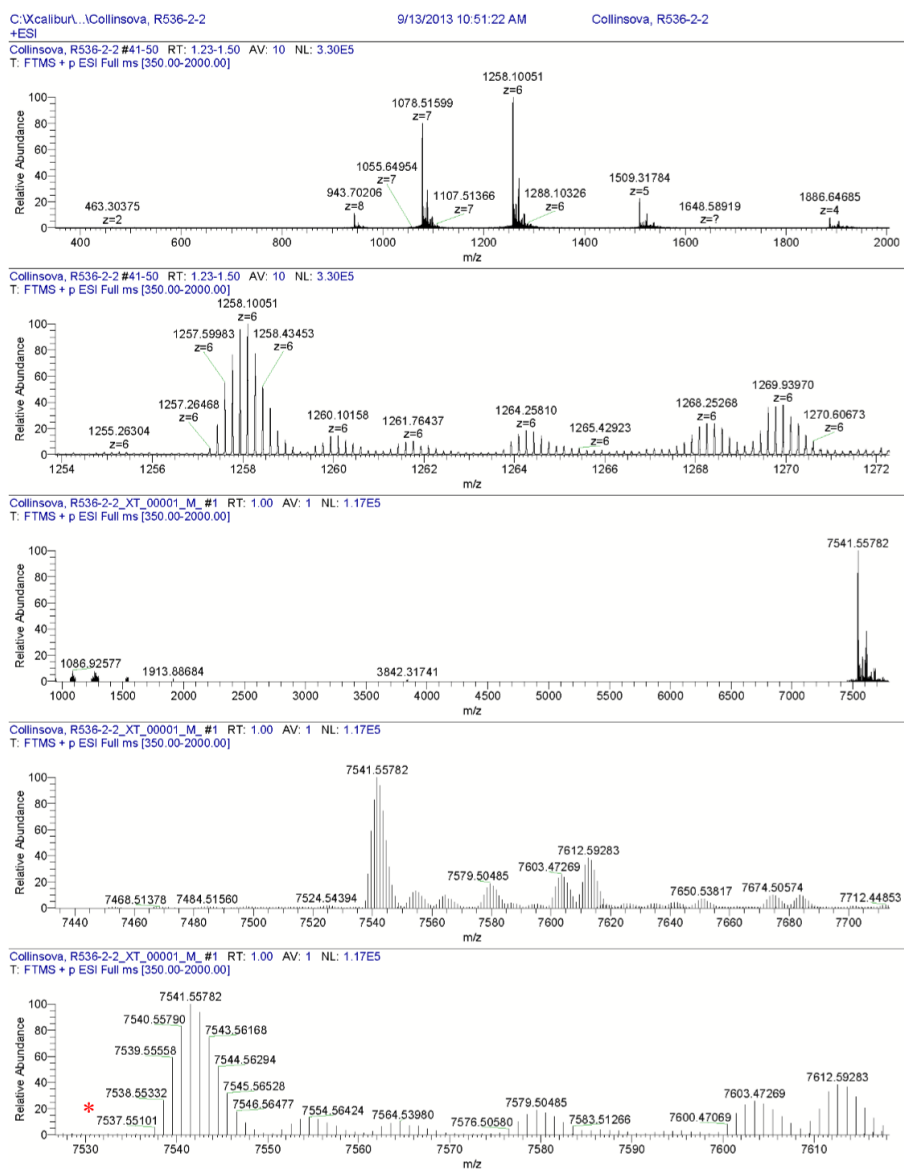


Figure S22. High-resolution ESI MS spectrum of purified analog **3**. Monoisotopic relative molecular weight (Mr) was calculated for $C_{328}H_{502}N_{92}O_{101}S_6$ and is 7537.530. The experimental monoisotopic Mr value was 7537.551 and the respective signal in the spectrum is marked by an asterisk.

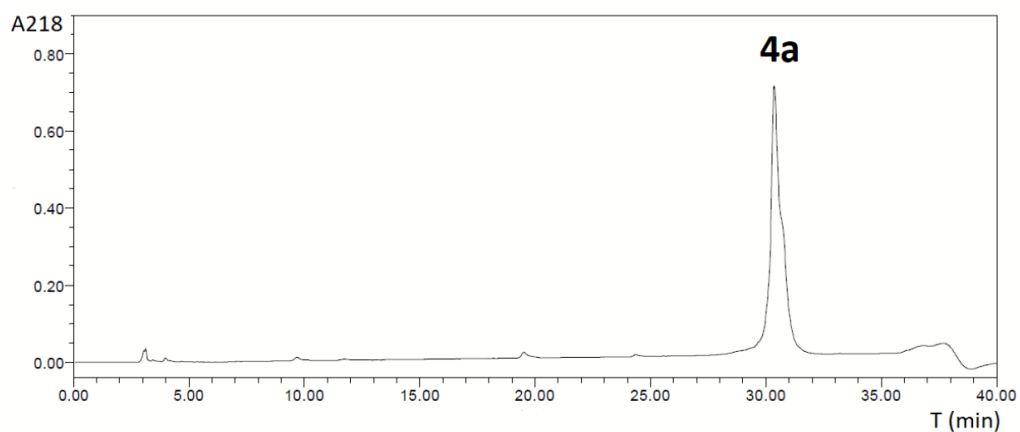


Figure S23. RP-HPLC analysis of the purified analog **4a**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/24 %, 30/40 %, 31/80 %, 32/80 %, 33/24%.

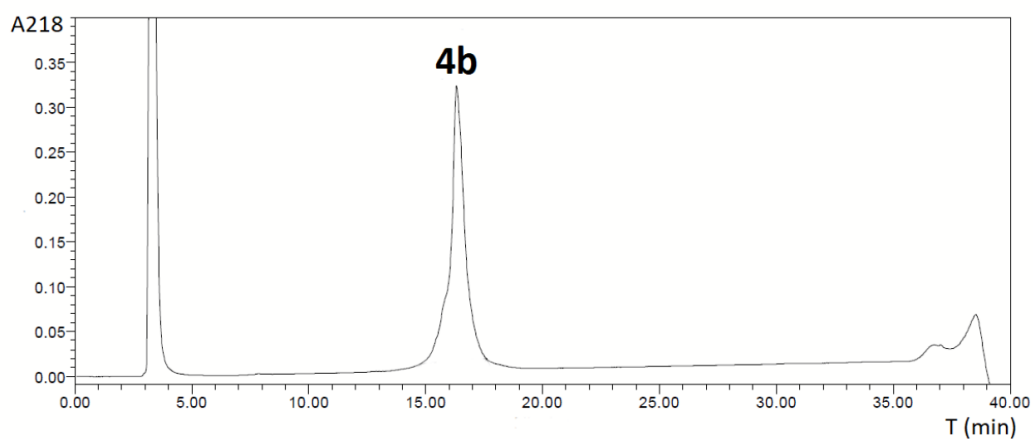


Figure S24. RP-HPLC analysis of the purified peptide **4b**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 30/40 %, 31/80 %, 32/80 %, 33/20 %. The peak at 4 min. is caused by elution of acetic acid, in which the analog was injected.

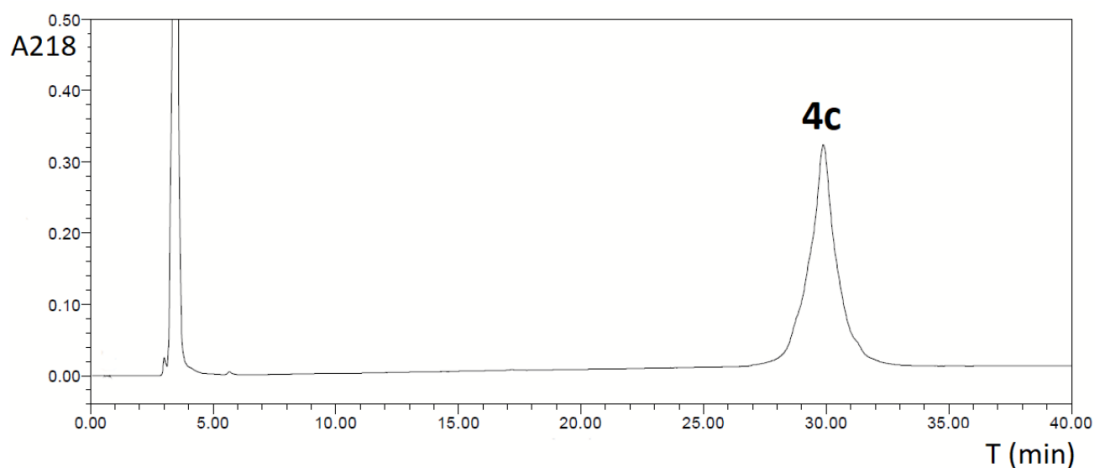


Figure S25. RP-HPLC analysis of the purified analog **4c**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 20/32 %, 40/35%, 41-42/80 %, 43/20 %. The peak at 4 min. is caused by elution of acetic acid, in which the analog was injected.

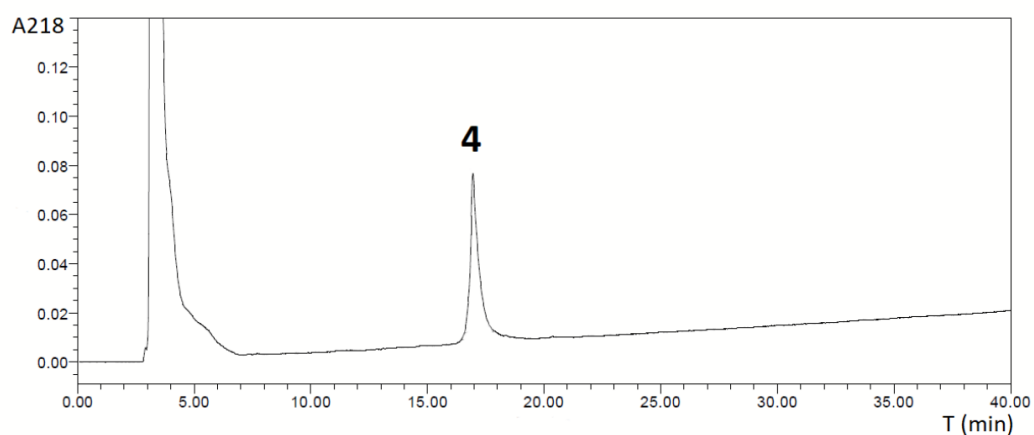


Figure S26. RP-HPLC analysis of the purified analog **4**. Vydac C4 column (214TP54, 5 μ m, 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 33/54 %, 34/80 %, 35/80 %, 36/20 %. The peak at 4 min. is caused by elution of acetic acid, in which the analog was injected. Native human IGF-1 analyzed under the same conditions is shown in Figure S17 B.

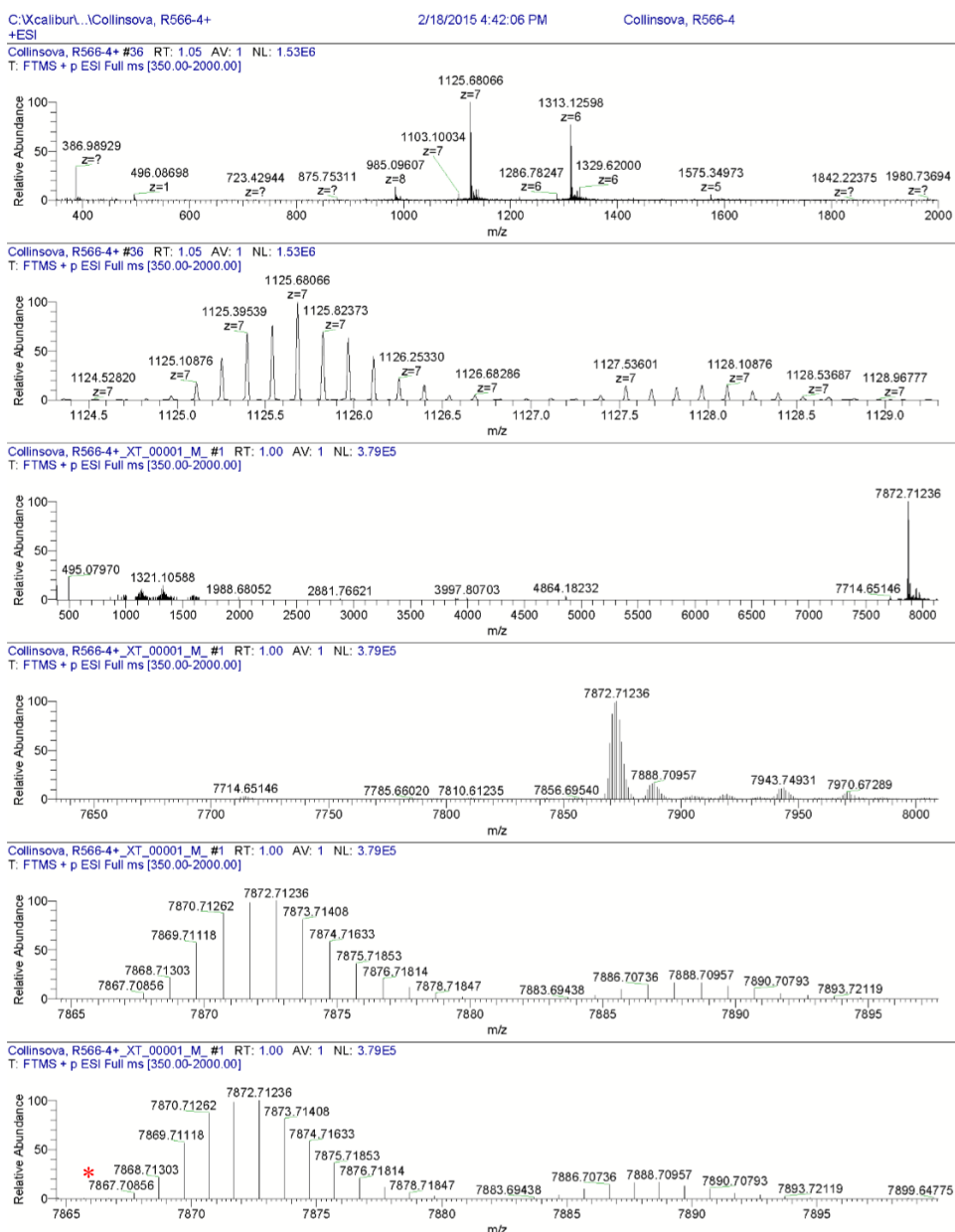


Figure S27. High-resolution ESI MS spectrum of purified analog **4**. Monoisotopic relative molecular weight (M_r) was calculated for $C_{339}H_{524}N_{100}O_{103}S_7$ and is 7867.688. The experimental monoisotopic M_r value was 7867.709 and the respective signal in the spectrum is marked by an asterisk.

Characterization of an aspartimide derivative of peptide 2a

MS analysis of the peptide found in peak **2ax** in the Figure S11 showed the major signal corresponding to the relative molecular mass smaller by 18 Da, compared to target peptide **2a** (see the main text). This difference indicated the loss of one water molecule and a possible formation of an aspartimide derivative. To confirm this hypothesis, the compound in peak **2ax** was isolated (Figure S28) and its structure analyzed by NMR.

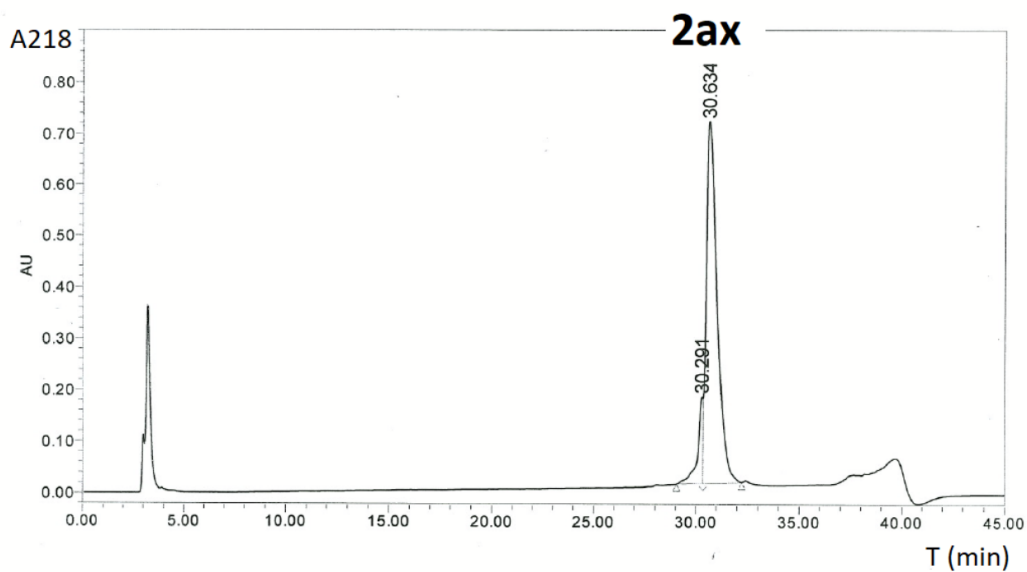


Figure S28. RP-HPLC analysis of the purified peptide **2ax**. Vydac C4 column (214TP54, 5 μm , 250 x 4.6 mm). Gradient (min/% ACN): 0/20 %, 33/40 %, 31/80 %, 32/80 %, 33/20 %.

NMR characterization of peptide **2ax**

NMR spectra were measured on a Bruker AVANCE-600 spectrometer (^1H at 600.13 MHz, ^{13}C at 150.9 MHz) equipped with cryoprobe in DMSO- d_6 solution at 310 K. The 2D-H,H-COSY, 2D-H,H-NOESY, 2D-H,C-HSQC and 2D-H,C-HMBC spectra were recorded and used for the structural assignment of proton and carbon signals.

The absence of the sequential NOE cross-peaks between Asp-20 and Arg-21 and amide NH peak of Arg-21 together with a doubling of the NH, H α and H β signals of Asp-20 indicated structural modification of the Asp20 residue, obviously to the racemic aspartimide. Aspartimide formation, which results from a ring-closure between β -COOH of Asp and amide NH of the following residue, is known as a frequently encountered side reaction in the solid-phase peptide synthesis (Behrendt, R.; White, P.; Offer, J. *Advances in Fmoc solid-phase peptide synthesis. Journal of Peptide Science* **2016**, *22*, 4-27). Proton NMR chemical shifts of peptide **2ax** are given in Table S1.

Table S1. Proton chemical shifts of peptide **2ax** in DMSO at 310 K.

Residue	NH	C α H	C β H	C γ H	C δ H	Other protons
Gly-1	n.a.	n.a.	--	--	--	--
Pro-2	--	4.398	2.089	1.938; 1.886	3.537; 3.477	--
Glu-3	8.344	4.275	1.950; 1.779	2.289; 2.225	--	
Thr-4	7.525	4.247	4.022	1.044	--	
Leu-5	7.895	4.321	1.619	1.498	0.872; 0.831	--
Cys(Acm)-6	8.064	4.417	2.912; 2.750	--	--	Acm: 4.251 (S-CH ₂); 8.452 NH); 1.85 (CH ₃)
Gly-7	8.115	3.727 (2H)	--	--	--	--
Ala-8	7.889	4.279	1.209	--	--	--
Glu-9	8.051	4.237	1.895; 1.768	2.233	--	
Leu-10	7.831	4.323			0.86; 0.82	
Val-11	7.694	4.129	1.960	1.032	--	--
Asp-12	8.136	4.545	2.679; 2.503	--	--	
Ala-13	7.773	4.199	1.036	--	--	--
Leu-14	7.780	4.115			0.848	--
Gln-15	7.810	4.139	1.808; 1.691	2.041	--	7.230 + 6.677 (CONH ₂)
Phe-16	7.771	4.452	3.082; 2.811	--	--	7.24 – 7.15 (C ₆ H ₅)
Val-17	7.850	4.533			--	--
Cys(Acm)-18	8.210	4.473	2.943; 2.761	--	--	Acm: 4.226 (S-CH ₂); 8.484 NH); 1.85 (CH ₃)
Gly-19	7.854	3.753	--	--	--	--
Asp-imid-20	8.618	4.593	2.967; 2.558	--	--	--
	8.586	4.417	2.995; 2.578	--	--	--
Arg-21	--	4.533	2.085; 1.870	1.535; 1.410	3.090	--
Gly-22	8.016	3.677 (2H)	--	--	--	--
Phe-23	7.618	4.450	2.892; 2.683	--	--	7.24 – 7.15 (C ₆ H ₅)
Tyr-24	7.974	4.387	2.849; 2.632	--	--	6.96 (H _{2,6}); 6.616 (H-3,5); 9.12 (OH)
Phe-25	8.012	4.548	3.047; 2.808	--	--	7.24 – 7.15 (C ₆ H ₅)
Asn-26	8.194	4.577	2.522; 2.437	--	--	7.310 + 6.842 (CONH ₂)
Lys-27	7.835	4.478	1.655; 1.579	1.466; 1.346	n.a.	
Pro-28	--	4.460	2.132; 2.037	1.887; 1.851	3.644; 3.505	--
Thr-29	7.749	4.152	4.002	1.040	--	
Gly-30	7.850	3.751; 3.661	--	--	--	--
Tyr-31	8.002	4.426	2.918; 2.670	--	--	7.013 (H _{2,6}); 6.636 (H-3,5); 9.12 (OH)
Gly-32	8.200	3.758 (2H)	--	--	--	--
Ser-33	7.878	4.386	3.623; 3.575	4.981 (OH)	--	--
Ser-34	8.012	4.362	3.677	5.025 (OH)	--	--
Ser-35	7.966	4.320	3.663; 3.580	5.058 (OH)	--	--
Ala(N3)-36	8.118	4.386	3.575	--	--	7.157 + 6.674 (CONH ₂)

Synthesis and characterization of a model heptapeptide and its aspartimide derivative

To further confirm the putative formation of an aspartimide derivative of the chain **2a**, we designed a model peptide Phe-Val-Cys(Acm)-Gly-Asp-Arg-Gly (**7** in Scheme S1), representing amino acids at the positions 16-22 of **2a**. The peptide **7** was prepared by the same solid-phase synthesis methodology on the Wang resin (ABI433A synthesizer) as **2a** (see in the main text). However, after the last Fmoc-protecting group was cleaved, the resin with a crude peptide was washed with DMF and the resin was suspended in a solution of 20 % piperidine in DMF and shaken at RT overnight. Then, the peptide was cleaved (95 % TFA, 2 % TIS, 2 % DODT and 1 % water for 2 hours and then precipitation by diethyl ether) and purified by a standard procedure. The HPLC chromatogram of the crude peptide **7** is shown in Figure S29. MS analysis of the compound in peak **7** identified peptide Phe-Val-Cys(Acm)-Gly-Asp-Arg-Gly (found Mr 823.3642, calculated 823.3647). MS analysis of the compound(s) in peak **7x** identified a compound with a relative molecular weight of 805.3562, which could indicate the presence of the formed aspartimide (calculated 805.3541), but also a minor signal of 823.3645, indicating the presence of racemized peptide **7**.

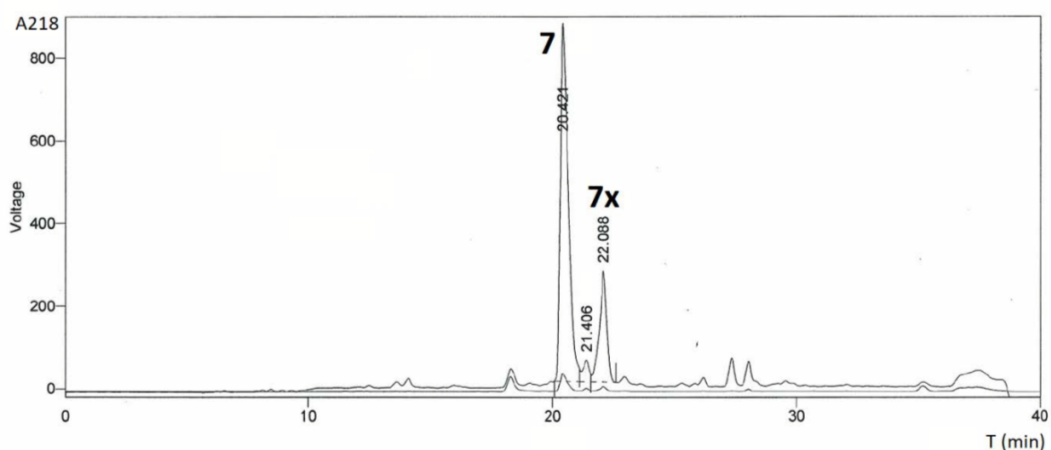


Figure S29. RP-HPLC analysis of the crude peptide **7**. Gradient (min/% ACN): 0 min /8 %, 30/40 %, 31/80 %. Macherey-Nagel C18 column (EC280/4 Nucleosil, 120-5, 250 x 4.6 mm).

Therefore, we further optimized the HPLC separation and found that peak **7x** can be divided into at least three separate peaks, which were labeled as **7xa**, **7xb** and **7xc** (Figure S30).

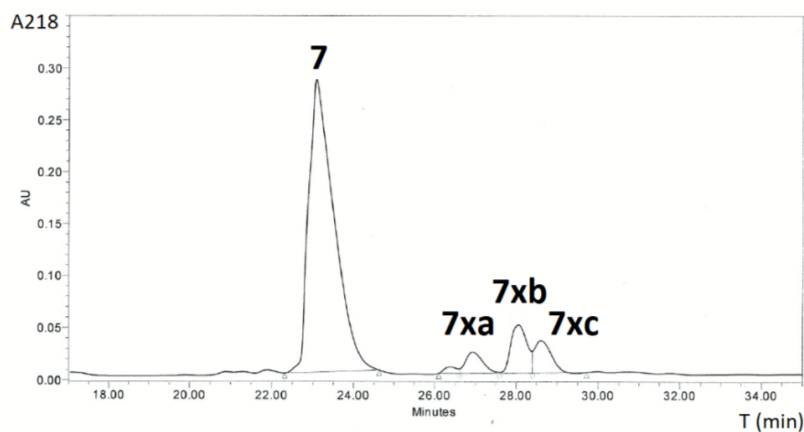


Figure S30. Optimized RP-HPLC analysis of the crude peptide **7**. Gradient (min/% ACN): 0 min/8 %, 30/16 %, 31/80 %, 32/8%. Macherey-Nagel C18 column (EC280/4 Nucleosil, 120-5, 250 x 4.6 mm).

Peaks **7**, **7xa**, **7xb** and **7xc** were separated and submitted for MS analysis, which revealed the presence of a compound having the molecular mass (823.3647 Da) of the target peptide Phe-Val-Cys(Acm)-Gly-Asp-Arg-Gly in the peaks **7** and **7xa**, but compounds with molecular masses 805.3537 and 805.3536 in peaks **7xb** and **7xc**, respectively. The NMR spectra confirmed structures of model peptides. Their ^1H and ^{13}C NMR data are summarized in Tables S2 and S3. In general, results of NMR analyses agree with MS data. Both pairs of peptides, i.e. **7-7xa** and **7xb-7xc**, differ obviously by the configuration at C_α carbon atom of Asp or Asp-imide, respectively. Although there is no direct evidence from NMR spectra, it is very probable that compounds in **7** and **7xb** have retained (*S*)-configuration, while compounds **7xa** and **7xc** contain (*R*)-configuration of Asp or Asp-imide, respectively. The deduced structures are shown in Scheme S1.

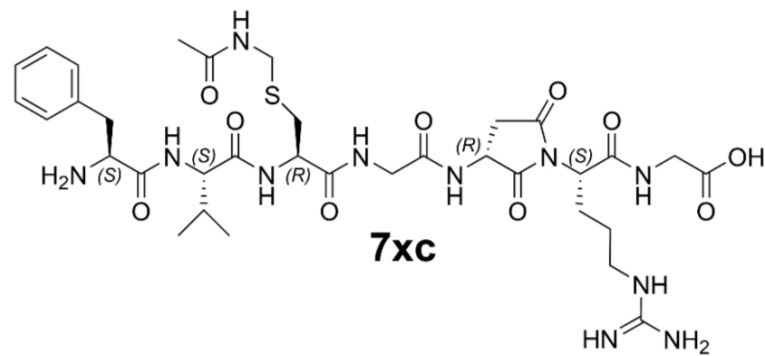
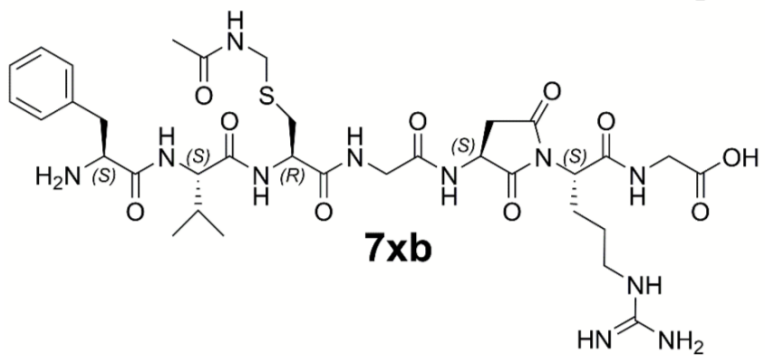
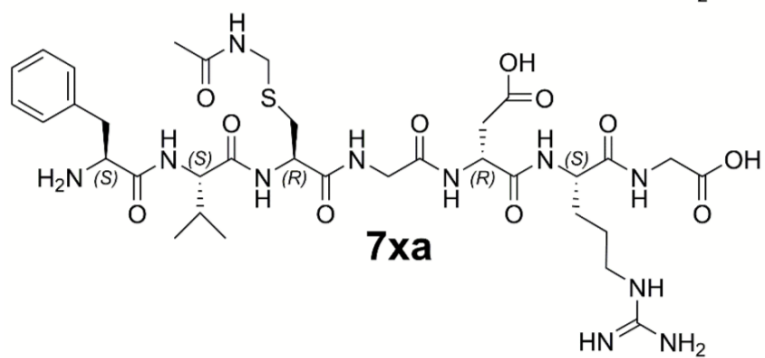
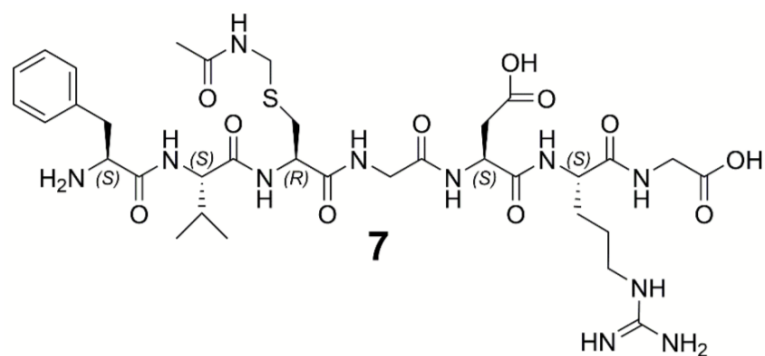
Table S2. Proton NMR data of peptides **7**, **7xa**, **7xb** and **7xc** in DMSO at 310 K. Coupling constants are given in brackets.

	7	7xa	7xb	7xc
Phe				
C α H	4.178 dd (7.5; 5.3)	4.166 dd (7.6; 5.0)	4.176 bt (7.8; 5.2)	4.173 bt (7.9; 5.3)
C β H	3.110 dd (14.4; 5.3)	3.102 dd (14.2; 5.0)	3.110 dd (14.2; 5.2)	3.108 dd (14.3; 5.3)
	2.946 dd (14.4; 7.5)	2.937 dd (14.2; 7.6)	2.940 dd (14.2; 7.8)	2.940 dd (14.3; 7.9)
C ₆ H ₅ : <i>o</i> -	7.25 m, 2H	7.23 m, 2H	7.25 m, 2H	7.25 m, 2H
<i>m</i> -	7.32 m, 2H	7.29 m, 2H	7.32 m, 2H	7.32 m, 2H
<i>p</i> -	7.26 m	7.25 m	7.27 m	7.27 m
Val				
NH	8.477 d (8.9)	8.467 d (9.0)	8.476 d (8.6)	8.467 d (8.7)
C α H	4.30 dd (8.9; 6.6)	4.334 dd (9.0; 6.4)	4.290 dd (8.6; 6.7)	4.300 dd (8.7; 6.7)
C β H	1.996 m	2.016 m	1.99 m	1.99 m
CH ₃	0.897 d (6.8)	0.918 d (6.8)	0.896 d (6.8)	0.900 d (6.8)
	0.878 d (6.8)	0.898 d (6.8)	0.879 d (6.8)	0.884 d (6.8)
Cys(Acm)				
NH	8.28 d (7.9)	8.370 d (8.3)	8.320 d (7.7)	8.348 d (7.5)
C α H	4.546 ddd (7.9; 4.9; 9.1)	4.607 ddd (8.3; 9.2; 4.6)	4.525 ddd (7.7; 5.0; 9.1)	4.520 ddd (7.7; 5.1; 9.0)
C β H	2.945 dd (13.8; 4.9)	2.975 dd (13.8; 4.6)	2.965 dd (13.8; 5.0)	2.952 dd (13.8; 5.1)
	2.756 dd (13.8; 9.1)	2.715 dd (13.8; 9.2)	2.765 dd (13.8; 9.1)	2.779 dd (13.8; 9.0)
S-CH ₂	4.286 dd (13.5; 6.3)	4.301 dd (13.5; 6.3)	4.281 dd (13.5; 6.4)	4.267 d, 2H (6.4)
	4.248 dd (13.5; 6.3)	4.184 dd (13.5; 6.3)	4.248 dd (13.5; 6.4)	
NH	8.493 t (6.3)	8.457 t (6.3)	8.493 t (6.4)	8.485 t (6.4)
CH ₃	1.855 s	1.853 s	1.854 s	1.856 s
Gly				
NH	8.11 t (5.8; 5.5)	8.170 t (5.8)	8.221 t (5.8)	8.257 t (5.8)
C α H	3.808 dd (16.8; 5.8)	3.837 dd (17.0; 5.8)	3.786 dd (16.8; 5.8)	3.796 dd (16.8; 5.8)
	3.734 dd (16.8; 5.5)	3.765 dd (17.0; 5.8)	3.755 dd (16.8; 5.8)	3.759 dd (16.8; 5.8)
Asp(Asp-imide)				
	Asp	Asp	Asp-imide	Asp-imide
NH	8.146 d (7.7)	8.196 d (7.6)	8.630 d (7.5)	8.627 d (7.6)
C α H	4.606 td (7.7; 7.8; 5.5)	4.609 td (7.6; 7.5; 5.8)	4.472 ddd (7.5; 9.4; 5.5)	4.595 ddd (7.6; 9.3; 5.4)
C β H	2.730 dd (16.7; 5.5)	2.710 dd (16.5; 5.8)	3.022 dd (17.6; 9.4)	2.966 dd (17.7; 9.3)
	2.522 dd (16.7; 7.8)	2.534 dd (16.5; 7.5)	2.576 dd (17.6; 5.5)	2.572 dd (17.6; 5.4)
COOH	12.42 br	12.00 br	--	--
Arg				
NH	7.95 d (8.1)	7.966 d (8.1)	--	--
C α H	4.272 dd (8.1; 7.8; 5.8)	4.269 m	4.515 dd (4.6; 10.6)	4.460 dd (4.7; 10.4)
C β H	1.74 m; 1.56 m	1.747 m, 1.564 m	2.12 m; 1.93 m	2.14 m; 1.93 m
C γ H	1.52 m (2H)	1.504 m, 2H	1.53 m; 1.44 m	1.52 m; 1.46 m
C δ H	3.093 m (2H)	3.09 m, 2H	3.074 m, 2H	3.080 m, 2H
NH	7.57 t (5.8)		7.538 t (5.7)	7.461 t (5.7)
Gly				
NH	8.085 t (5.9; 5.9)	8.077 t (5.8)	8.132 t (6.3; 5.5)	8.108 t (6.3; 5.5)
C α H	3.777 dd (17.5; 5.9)	3.775 dd (17.5; 5.8)	3.840 dd (17.4; 6.3)	3.847 dd (17.4; 6.3)
	3.740 dd (17.5; 5.9)	3.737 dd (17.5; 5.8)	3.595 dd (17.4; 5.5)	3.585 dd (17.4; 5.5)

Table S3. Carbon-13 chemical shifts of peptides **7**, **7xa**, **7xb** and **7xc** in DMSO at 310 K.

	7	7xa	7xb	7xc
Phe				
C=O	167.93	168.06	167.94	167.93
C α	53.21	53.29	53.20	53.19
C β	37.10	37.13	37.09	37.10
C ₆ H ₅ <i>ipso-</i>	134.83	134.84	134.81	134.78
<i>ortho-</i>	129.63	129.59	129.63	129.62
<i>meta-</i>	128.60	128.58	128.60	128.60
<i>para-</i>	127.18	127.15	127.18	127.18
Val				
C=O	170.58	170.60	170.66	170.73
C α	57.91	57.97	57.91	57.84
C β	30.95	30.94	30.93	30.97
CH ₃	19.26	19.34	19.19	19.18
	18.24	18.14	18.26	18.25
Cys(Acm)				
C=O	170.47	170.49	170.46	170.51
C α	52.87	52.23	52.97	53.08
C β	32.68	33.30	32.54	32.47
S-CH ₂	40.50	40.40	40.53	40.54
Ac: C=O	169.88	169.73	169.89	169.87
CH ₃	22.64	22.63	22.64	22.63
Gly				
C=O	168.25	168.70	169.51	169.52
C α	42.09	42.06	41.96	42.00
Asp(imide)				
C=O	171.83	171.85	175.25	175.78
C α	49.62	49.67	48.44	48.63
C β	36.31	36.35	35.00	34.82
COOH	170.48	170.52	--	--
C=O	--	--	174.75	174.32
Arg				
C=O	171.52	171.48	168.09	168.00
C α	52.22	52.23	52.80	53.15
C β	29.07	29.03	24.51	24.63
C γ	24.89	24.83	25.16	25.51
C δ	40.52	40.52	40.42	40.41
NH=C-NH ₂	156.96	156.93	156.93	156.90
Gly				
COOH	170.94	170.95	170.88	170.88
C α	40.73	40.75	41.00	40.99

Scheme S1. Deduced structures of compounds **7**, **7xa**, **7xb** and **7xc**



7.2 Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides

Background:

The rise of Cu^I-catalyzed click chemistry lead to an increased demand for azido derivatives of amino acid as precursors for this synthesis. Some azido amino acids are commercially available, however their costs are high. Thus, we investigated the possibility of the in-house synthesis of five Fmoc azido amino acids: β -azido L-alanine and D-alanine, γ -azido L-homoalanine, δ -azido L-ornithine and ω -azido L-lysine.

Summary:

We investigated several reaction pathways for the synthesis of several N-Fmoc-protected azido amino acids described in the literature and we suggested several improvements. Moreover, we proposed several alternative routes for the synthesis of these compounds in a high purity. In conclusion, we provided detailed synthetic protocols for multigram preparation of these Fmoc azido amino acids, which can be prepared in a week or two and with user-friendly costs.

My contribution:

Under the guidance of Dr. Jan Pícha, I prepared a Fmoc- β -azido-L-alanine which I used for synthesis of IGF-1 analogues and I participated on synthesis of model peptides with azido amino acids in the sequence.



Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides

Jan Pícha, Miloš Buděšínský, Kateřina Macháčková, Michaela Collinsová and Jiří Jiráček*

The rise of CuI-catalyzed click chemistry has initiated an increased demand for azido and alkyne derivatives of amino acid as precursors for the synthesis of clicked peptides. However, the use of azido and alkyne amino acids in peptide chemistry is complicated by their high cost. For this reason, we investigated the possibility of the in-house preparation of a set of five Fmoc azido amino acids: β -azido L-alanine and D-alanine, γ -azido L-homoalanine, δ -azido L-ornithine and ω -azido L-lysine. We investigated several reaction pathways described in the literature, suggested several improvements and proposed several alternative routes for the synthesis of these compounds in high purity. Here, we demonstrate that multigram quantities of these Fmoc azido amino acids can be prepared within a week or two and at user-friendly costs. We also incorporated these azido amino acids into several model tripeptides, and we observed the formation of a new elimination product of the azido moiety upon conditions of prolonged couplings with 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate/DIPEA. We hope that our detailed synthetic protocols will inspire some peptide chemists to prepare these Fmoc azido acids in their laboratories and will assist them in avoiding the too extensive costs of azidopeptide syntheses.

Experimental procedures and/or analytical data for compounds 3–5, 20, 25, 26, 30 and 43–47 are provided in the supporting information. © 2017 The Authors Journal of Peptide Science published by European Peptide Society and John Wiley & Sons Ltd. Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: synthesis; azido amino acid; alanine; homoalanine; ornithine; lysine; azide elimination

Introduction

The discovery of the CuI-catalyzed Huisgen 1,3-dipolar cycloaddition reaction of azides and alkynes, which leads to the formation of 1,4-disubstituted 1,2,3-triazoles [1,2], has triggered a boom of these types of click reactions, which have subsequently found a plethora of applications in peptide and protein chemistry. It is because 1,2,3-triazoles present a motif with structural and electronic characteristics similar to those of the peptide bond [3,4]. 1,2,3-Triazoles have also found applications in cyclizations [5,6], induction of β -turns [7,8], β -hairpins [9], helical structures [10,11] or mimics of disulfide bonds [12] in peptides. The broad applications of 1,2,3-triazoles were reviewed in many excellent reviews (e.g. [13,14] or [15]). Moreover, azides have also found utility in a variant of Staudinger ligation for the synthesis of peptides and proteins [16]. However, the syntheses of peptides with azido or alkyne moieties are often hampered by the high cost of azido or alkyne precursors, mostly Fmoc-protected azido and alkyne amino acids, which are usually available for about €250–300 per 250 mg. This can make the preparation of larger series of azido/alkyne peptides very expensive, as we have recently experienced [17]. For this reason, we decided to investigate the accessibility of the in-house preparation of a series of five Fmoc-protected azido amino acids; (S)-2-amino-3-azidopropanoic and (R)-2-amino-3-azidopropanoic acids (β -azido L-alanine and D-alanine), (S)-2-amino-4-azidobutanoic acid (γ -azido L-homoalanine), (S)-2-amino-5-azidopentanoic acid (δ -azido L-ornithine) and (S)-2-amino-6-azidohexanoic acid (ω -azido L-lysine).

Serine is a convenient precursor for the synthesis of the azido derivative of alanine. Several different methods were previously

developed for introducing the azido moiety to a serine derivative: (i) ring opening of cyclic *N*-(phenyl fluoride) serine sulfamidate [18–20], (ii) opening of (S)-3-amino-2-oxetanone [21], (iii) mild bromination of the hydroxyl group followed by azidation [22–25], (iv) Mitsunobu reaction treatment with triphenylphosphine (PPh₃), hydrazoic acid and azodicarboxylate [26–35], and (v) activation of the hydroxyl group by mesyl chloride (MsCl) [35–38] or tosyl chloride [39], followed by substitution with sodium azide. In addition, the synthesis of azidoalanine starting from *N*-protected asparagine represents a different but straightforward approach. The first step is Hoffman degradation by treatment with *I*-bis(trifluoroacetate) (CAS 2712-78-9) [40–44] or (diacetoxyiodo)benzene (PIDA) [45], followed by a diazotransfer reaction [41,46–51].

* Correspondence to: Jiří Jiráček, Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, v.v.i., Flemingovo nám. 2, 166 10 Prague 6, Czech Republic. E-mail: jiracek@uochb.cas.cz

Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, v.v.i., Flemingovo nám. 2, 166 10, Prague 6, Czech Republic

Abbreviations: Bn, benzyl; Boc₂O di-tert-butyl dicarbonate, CAS 24424-99-5; Fmoc-OSu *N*-(9-fluorenylmethoxycarbonyloxy)succinimide, CAS 82911-69-1; MsCl mesyl chloride, CAS 124-63-0; PIDA (diacetoxyiodo)benzene, CAS 3240-34-4; TfN₃ triflic azide, CAS 3855-45-6; triflic anhydride trifluoromethanesulfonic anhydride, CAS 358-23-6; Z, benzyloxy carbonyl.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

In general, there are two main approaches for the synthesis of derivatives of Fmoc-L-azidohomoalanine. Mostly, the protected Fmoc-L-glutamine is converted under Hofmann rearrangement conditions [41,52–54] to 2-Fmoc-4-aminobutanoic acid, followed by the azido transfer reaction [12,41,46]. The other approach involves protected L-aspartic acid, which is partially reduced via a mixed anhydride. The resulting alcohol is mesylated, and the corresponding mesyl derivative is replaced by the azido function [55].

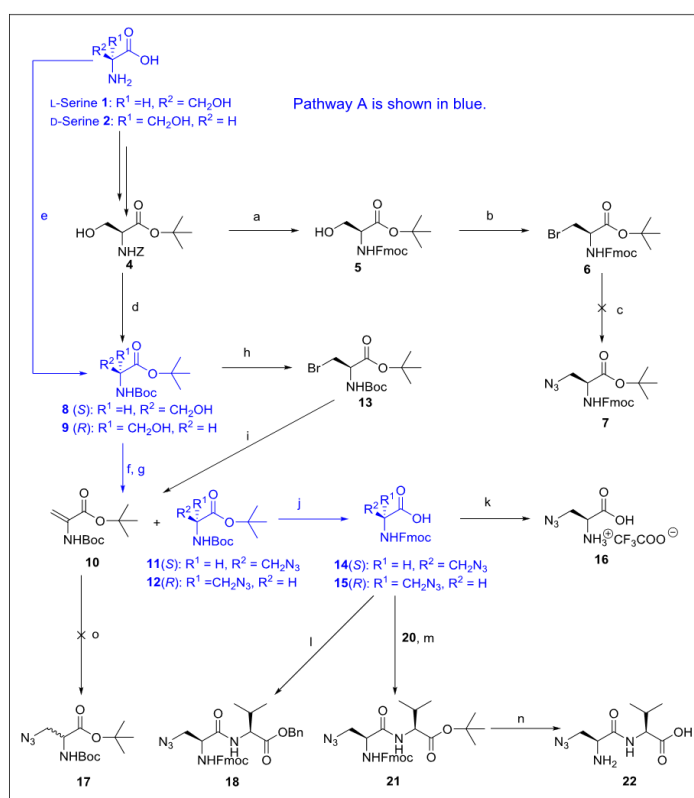
The strategy of orthogonal functional group protection [56–59] of L-ornithine and L-lysine is usually used for the synthesis of their azido derivatives.

Here, we compared several of these reaction pathways and also investigated a few new routes for the preparation of the above-listed Fmoc-protected azido amino acids in a multigram scale, including their incorporation into short model peptides. The advantages and drawbacks of the approaches are discussed.

Results and Discussion

Synthesis of β -Azido L-Alanine and D-Alanine

Firstly, L-serine **1** was chemoselectively protected by the benzyloxycarbonyl (Z) group, and Z-Ser **3** was then esterified with *tert*-butyl bromide in the presence of a benzyltriethylammonium chloride phase catalyst and excess of potassium carbonate [60,61] (Scheme 1). These two transformations led to the intermediate **4**. The removal of the Z group by a catalytic hydrogenation and a subsequent acylation with *N*-(9-fluorenylmethoxycarbonyloxy) succinimide (Fmoc-OSu) yielded Fmoc-protected compound **5**, which, after the Abell reaction conditions, gave compound **6**. It is well known that the direct azidation of bromo compounds in DMF (CAS 68-12-2) is unfeasible because of the high basicity of NaN_3 . Therefore, for the introduction of N_3 moiety under mild



Scheme 1. Reagents, conditions and yields: (a) 10% Pd/C, H_2 , methanol, RT, overnight then Fmoc-OSu, NaHCO_3 , water and dioxane, 0°C for 1 h, then RT overnight (74% over two steps); (b) CBr_4 , PPh_3 , DCM 0°C for 1 h, then RT overnight (84%); (c) Trimethylsilyl azide, hexamethylphosphoramide, 70°C overnight; (d) 10% Pd/C, H_2 , methanol, RT overnight, then Boc_2O , NaHCO_3 , water and dioxane, 0°C for 1 h, followed at RT overnight, then $\text{C}_4\text{H}_9\text{Br}$, K_2CO_3 , benzyl triethylammonium chloride, *N,N*-dimethyl acetamide, 60°C overnight (72% over two steps for **8**, 67% over two steps for **9**); (f) MsCl , TEA, DCM, 0°C for 0.5 h, then NaN_3 , DMSO, 70°C overnight (16% for **10**, 47% for **11** over two steps, 21% for **10**, 50% for **12** over two steps); (g) MsCl , pyridine, DCM, 0°C for 0.5 h, then NaN_3 , DMSO, 70°C overnight (26% for **10**, 46% for **11** over two steps); (h) CBr_4 , PPh_3 , DCM, 0°C for 1 h, then RT overnight (68%); (i) NaN_3 , DMSO, 70°C overnight (40% for **10**, 41% for **11** over two steps); (j) TFA, DCM, water, RT overnight, then NaHCO_3 and Fmoc-OSu, dioxane and water, 0°C for 1 h, then RT overnight (86% for **14** over two steps, 86% for **15** over two steps); (k) 2-chlorotrityl chloride resin, DIPEA, DMF, 2 h, then piperidine, DMF (75%); (l) L-Val-OBn-TsOH, DIPEA, PyBroP, DMF RT overnight (74%); (m) **20**, DIPEA, PyBroP, DMF, RT overnight (70%); (n) TFA, DCM, 2 h, 2-chlorotrityl chloride resin, DIPEA, DMF, 2 h, then piperidine, DMF (67%); (o) NaN_3 , DMSO, 70°C overnight.

conditions, we decided to apply trimethylsilyl azide (CAS 4648-54-8) in hexamethylphosphoramide (CAS 680-31-9) or DMF [62,63]. However, we surprisingly only isolated from the reaction mixture starting compound **6** instead of **7**. The failure of this synthetic pathway indicated that the sequence of steps must be changed; the azido moiety should be introduced first, followed by the deprotection of the amino group and then by derivatization with Fmoc. Thus, the Z group was again subjected to catalytic hydrogenolysis, and then the free amine group was protected by Boc to afford derivative **8**. We also found that the key intermediate **8** can be achieved by a shortcut (Scheme 1, pathway A) directly from L-serine **1**. Thus, the Boc-L-Ser-OTBu **8** was prepared in a good yield (72%) by the reaction of L-serine with *tert*-BuBr in a similar manner as compound **4**. Thereafter, activation of the hydroxyl group was performed with MsCl under basic conditions (TEA, CAS 121-44-8), ensuring the stability of acid-sensitive groups. The reaction mixture was tested with TLC, and except for the required mesyl intermediate **11**, the dehydroalanine **10** was also observed as the elimination product [31,64]. We also tested pyridine as a weaker base instead of TEA, but the full conversion of **8** to mesyl intermediate **11** was achieved only after 24 h, but again accompanied by a significant formation of **10**. To prevent this unwanted process, compound **8** was transformed to bromide derivative **13** and subsequently reacted with sodium azide in DMSO (CAS 67-68-5). In this case, **11** was obtained in a similar yield, but with a substantially higher amount of the elimination product **10**. Probably, a more easily leaving Br⁻ group better promotes the elimination process. Hence, we can conclude that both mesylation and azidation reactions contribute to the formation of dehydroalanine **10**. Finally, with the use of standard procedures, both acid-labile protecting groups of **11** were removed with TFA (CAS 76-05-1) and the amino group was acylated with Fmoc-Osu to furnish the required product **14**.

Next, starting from D-serine **2** and through intermediates **9** and **12**, we employed the synthetic pathway A and the following reactions for the preparation of Fmoc-β-azido-D-Ala **15**.

Cumulative yields of the six-step synthetic pathway A (i.e., (i) acylation with Boc, (ii) alkylation with *tert*-But, (iii) activation by Ms, (iv) addition of azide, (v) elimination of acid-labile groups and (vi) acylation with Fmoc) were satisfactory: 29% for **14** and 28% for **15**, both in multigram scales.

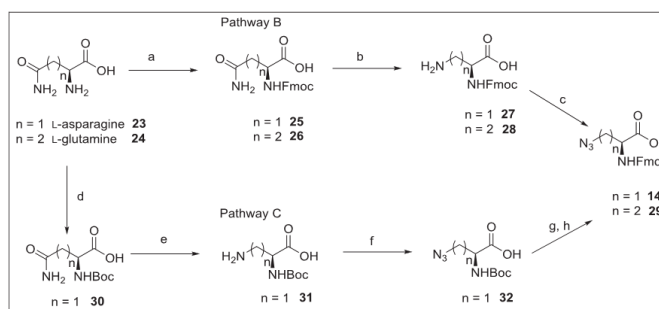
Johansson and Pedersen [65] and others [35,66] claimed that dehydroalanine **10** is a perfect Michael acceptor, which undergoes

racemization. To verify this statement, we carried out a simple experiment; the isolated dehydroalanine **10** was heated in DMSO with an excess of NaN₃ at 70 °C overnight. However, no traces of the expected product **17** were found.

Next, we verified the optical purity of the compound **14** by NMR spectroscopy. For this, the dipeptide **18** was prepared by the reaction of **14** with L-Val-OBn-TsOH. The incorporation of valine added a new stereo center to the molecule. However, the presence of the sterically bulky Fmoc group resulted in the observation of geometrical isomers of the carbamate in a ratio of 60 : 40. Therefore, we had to prepare a fully deprotected molecule. Compound **14** was thus coupled with ester **20**, which was prepared by esterification of L-valine **19** using *tert*-butyl acetate [67] as a source of (CH₃)₃C⁺ cation. Acidic hydrolysis gave free acid, which was attached to 2-chlorotrityl chloride resin. This allowed the removal of the Fmoc protecting group by conveniently washing off poorly separable dibenzofulvene. A usual work-up and separation furnished dipeptide **22**, which was manifested by only one set of signals in both ¹H and ¹³C NMR spectra. The same protocol with the chlorotrityl resin was used for the preparation of free amino acid **16**. The optical purity of water-soluble acid **16** was checked by the method of Inamoto *et al.* [68]. No splitting of signals on α-carbons and α-protons was observed after the addition of sodium [(S)-1,2-diaminopropane-*N,N,N',N'*-tetraacetato]-samarate(III) [69] to a pH-adjusted solution of **16** in a 2 : 1 molar ratio. In conclusion, NMR analyses unequivocally proved the high optical purity of compound **14**. We have not performed the same procedure with the optical isomer **15**, which was prepared by the same reactions, but from the pure D-serine. However, the same high optical purity can be expected.

We also investigated two alternative reaction routes (Scheme 2) for the synthesis of Fmoc-β-azido-L-Ala **14** as outlined in Scheme 2. Firstly, L-asparagine **23** was protected by the Fmoc group; then carboxamide functionality of the resulting intermediate **25** was eliminated under Hoffmann rearrangement conditions. The last step of this reaction pathway B was a diazotransfer reaction, which allowed the conversion of the amino group to the corresponding azido acid **14**. In parallel, we demonstrated that **14** can be obtained with a similar synthetic strategy, but using Boc protection and starting from L-asparagine **23** over intermediates **31** and **32** (reaction pathway C).

Compound **14**, which was prepared using three different synthetic pathways (A, B or C), provided the same mass and NMR



Scheme 2. Reagents, conditions and yields: (a) Na₂CO₃, Fmoc-OSu, dioxane and water, 0 °C for 1 h, then RT overnight (87% for **25**, 93% for **26**); (b) PhI(OAc)₂, CH₃CN, ethyl acetate and water at RT overnight (75% for **27**, 56% for **28**); (c) TfN₃, NaHCO₃, CuSO₄·5H₂O, water and methanol at RT overnight (80% for **14**, 92% for **29**); (d) Na₂CO₃, Boc₂O, dioxane and water 0 °C 1 h then RT overnight (73%); (e) PhI(OAc)₂, CH₃CN, ethyl acetate and water at RT overnight (75%); (f) TfN₃, TEA, CuSO₄·5H₂O, water and methanol at RT overnight; (g) TFA, DCM, 2 h at RT; (h) NaHCO₃, Fmoc-OSu, dioxane and water, 0 °C for 1 h, then RT overnight (37% over three steps).

spectra and other physicochemical characteristics. Therefore, the number of synthetic steps, cumulative yields and costs are decisive for the choice of the optimal strategy. Pathway A includes six steps and gave 29% yield, pathway B was performed in three steps and with 52% yield and pathway C required five steps and gave 20% yield. Clearly, from this aspect, the preferred synthetic route is pathway B and the less convenient is pathway C. We also calculated the approximate costs of synthetic pathways A and B for the preparation of **14**, using precursor and solvents purchased at standard prices from Fluka. Synthetic pathway A yielded 1 g of **14** for about €37 and pathway B for about €43. Pathway A can be completed within a week; pathway B is faster. Taking everything together, the method of choice for the preparation of **14** is pathway B, despite the fact that its diazotransfer reaction requires the use of an excess of rather costly trifluoromethanesulfonic anhydride (triflic anhydride).

Synthesis of β -Azido L-Homoalanine

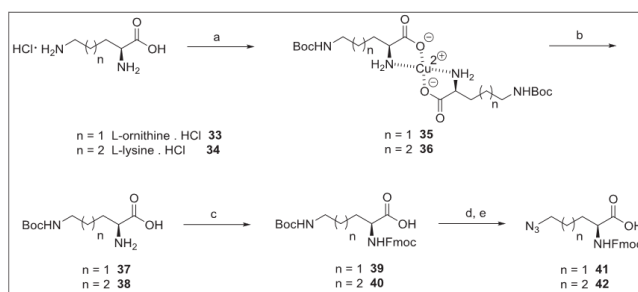
Next, for the preparation of L-homoazidoalanine, we chose the straightforward pathway B (Scheme 2), starting from L-glutamine over intermediates **26** and **28**. The product **29** was obtained in an excellent yield of 92%.

Synthesis of δ -Azido L-Ornithine and ω -Azido L-Lysine

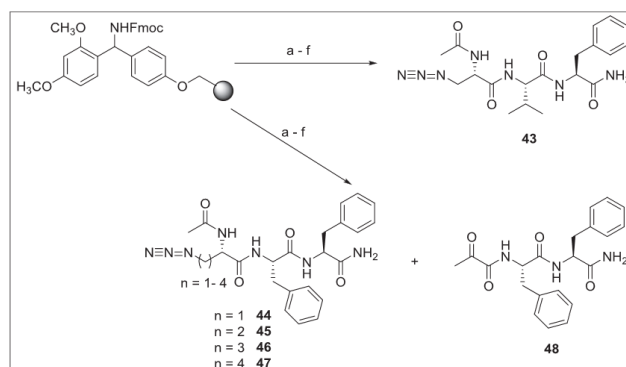
For the synthesis of azido acids **41** and **42**, derivatives of L-ornithine and L-lysine, we applied the strategy of orthogonal functional group protection [56–59] (Scheme 3). Synthesis started from commercially available L-ornithine-HCl **33** or L-lysine-HCl **34**, and the reaction with copper acetate monohydrate under basic conditions afforded [Orn(Boc)]₂Cu **35** or [Lys(Boc)]₂Cu **36**, respectively, which were isolated by a perfect filtering off of their insoluble copper complexes. Metal was quantitatively removed using 8-quinolinol to furnish selectively Boc-protected intermediates **37** and **38** in forms of zwitterions. The alpha amino group was acylated with Fmoc-OSu, and the resulting diamino acids **39** and **40** were treated with TFA to liberate δ -free amine and ω -free amine, respectively. The final step is represented by a diazotransfer reaction, which leads to the required Fmoc-azido-L-norvaline **41** or Fmoc-azido-L-lysine **42**.

Synthesis of Model Azido Tripeptides

Finally, we synthesized a series of model tripeptides **43–47**, using the standard manual Fmoc solid-phase synthesis protocol [70] (Scheme 4). When the methodology was precisely followed, all required peptides were obtained in good yields and with a high



Scheme 3. Reagents, conditions and yields: (a) CuAc₂·H₂O, water, then Boc₂O in acetone overnight (90% for **35**, 92% for **36**); (b) 8-hydroxyquinoline, water, 4 h (90% for **37**, 88% for **38**); (c) Fmoc-OSu, NaHCO₃, water and dioxane, 0 °C for 0.5 h, then RT overnight (93% for **39**, 95% for **40**); (d) TFA, DCM, 2 h at RT (e) TfN₃, NaHCO₃, CuSO₄·5H₂O, water and methanol at RT overnight (92% for **41** over two steps, 89% for **42** over two steps).



Scheme 4. Reagents, conditions and yields: (a) 20% piperidine/DMF; (b) (i) Fmoc-L-Phe, HBTU, DIPEA, DMF, 2 × 2 h at RT (ii) 20% piperidine/DMF; (c) (i) Fmoc-L-Phe or Fmoc-L-Val, HBTU, DIPEA, DMF, 2 × 2 h at RT (ii) 20% piperidine/DMF; (d) **14**, **29**, **41** or **42**, HBTU, DIPEA, DMF, 2 × 2 h at RT (ii) 20% piperidine/DMF; (e) Ac₂O, DIPEA, DMF, 15 min at RT; (f) 95% TFA/water, 1 h at RT. Yields (after HPLC purification) 69% for **43**, 76% for **44**, 80% for **45**, 77% for **46** and 71% for **47**. In the case of long couplings (5 and 18 h) with **14**, the yield for **44** was only 17% accompanied by the presence of **48** in 32% yield.

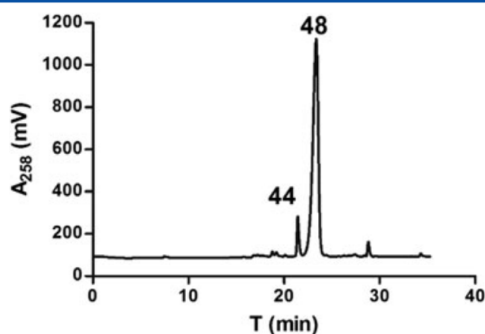


Figure 1. HPLC trace of the purification of the crude tripeptide **44**. The major isolated products (**44** and **48**) are labeled with compound numbers.

chemical purity. However, surprisingly, during the synthesis of **44** and only in the case of prolonged condensations (5 and 18 h) of Fmoc-azidoalanine **14** with the resin-bound Phe-Phe-NH₂, we observed the massive appearance of a new compound representing a major product of the synthesis (Figure 1). This product was isolated, and its chemical structure assigned using spectral methods and attributed to the compound **48**. It appears that only alpha-keto functionality is present in **48** instead of methylazido and acetylamino moieties. We have not found any information in the available literature about this type of elimination of azido species. Some studies (e.g., [71]) reported on obtaining carbonyl compounds from α,β -disubstituted compounds. However, here, we can only speculate that the shorter side chain of **14** (i.e., the proximity of the azido moiety and the primary α -amino group) and/or possibly also the activating agent 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) in the presence of DIPEA (7087-68-5) and/or longer couplings or treatment with 95% TFA/5% water may play a role in this side reaction. Interestingly, in addition, it seems that this elimination is also sequence specific, as it occurred only in the case of the coupling of **14** to Phe-Phe dipeptide and not to other sequences. Hence, at this stage, we rather do not suggest any plausible reaction mechanism for this process. The susceptibility of **14** to the elimination of its azido moiety during the synthesis of a simple tripeptide suggests that the use of azido amino acids in the peptide synthesis is not without risks and that some precautions should be taken (e.g., shorter reaction times and alternative reagents).

Conclusions

In conclusion, we investigated several synthetic protocols for the preparation of L-Fmoc- β -azidoalanine and D-Fmoc- β -azidoalanine (**14** and **15**, respectively). We found that pathway B starting from asparagine is the most straightforward one and it can also be used for the preparation of γ -azido L-homoalanine when starting from glutamine. NMR analysis confirmed the high optical purity of **14** prepared with these protocols. We also synthesized L-Fmoc- γ -azidohomoalanine **29**, L-Fmoc- δ -azidoornithine **41** and L-Fmoc- ϵ -azidolysine **42**. Several synthetic steps previously described in literature were improved and optimized, and several new reactions investigated. All synthetic procedures were described in detail, and the complete physicochemical characterization of all intermediates and final compounds was provided. We found that multigram

quantities of these Fmoc-protected azido amino acids can be prepared within a week, at average costs of about €40 per gram of final compounds (excluding work and energies). This makes them incomparably cheaper than standard commercial counterparts. We also observed a new type of elimination which occurred during prolonged couplings upon the solid-phase synthesis of Ac- β -azido-Ala-Phe-Phe-NH₂.

Experimental Part

Reagents and solvents (Sigma-Aldrich-Fluka, St. Louis, MO, USA) used in this study were of analytical grade. TLC analyses were performed on silica-gel-coated aluminum plates (Fluka). The compounds were visualized by exposure to UV light at 254 nm and by ninhydrin spraying (dark blue color) of Boc-protected amines or amines. Flash chromatography purifications were carried out on silica gel (40–63 μ m, Fluka). Preparative RP-HPLC chromatography was carried out on a C18 Luna column (Phenomenex, Torrance, CA, USA, 250 \times 21.2 mm, 10 μ m) at a flow rate 9 ml/min (solvent A: 0.1% TFA; solvent B: 80% CH₃CN, 0.1% TFA). Eluted compounds were detected at 218 and 254 nm and lyophilized from water. Melting points were determined on a Boetius block and are uncorrected. ¹H and ¹³C NMR spectra were measured on a Bruker AVANCE-600 spectrometer (Billerica, MA, USA; ¹H at 600.13 MHz, ¹³C at 150.9 MHz) in CDCl₃, DMSO-*d*₆, CD₃OD or D₂O solution at 300 K. The 2D-H,H-COSY, 2D-H,C-HSQC and 2D-H,C-HMBC spectra were recorded and used for the structural assignment of proton and carbon signals. IR spectra were recorded on Bruker IFS 55 Equinox apparatus. HRMS spectra were obtained on an FTMS mass spectrometer LTQ-orbitrap XL (Thermo Fisher, Bremen, Germany) in electrospray ionization mode.

Experimental procedures and analytical data for compounds **3–5** are provided in the supporting information.

tert-Butyl 2-(5)-(9-fluorenylmethyloxycarbonylamino)-3-bromopropanoate **6**

Ester **5** (14.6 g; 38.1 mmol) and CBr₄ (15.2 g; 45.7 mmol) were dissolved in 100 ml of DCM (CAS 75-09-2). The flask with the reaction mixture was immersed in an ice cooling bath, and PPh₃ (12 g; 45.7 mmol) in 100 ml DCM was added dropwise under stirring. Stirring at 0 °C was continued for 1 h and then at room temperature (RT) overnight. DCM was removed under reduced pressure, and a brown residue was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in petroleum ether. The product was a colorless oil, which was triturated in petroleum ether at –20 °C. Yield 14.2 g (84%). White solid. m.p. 75–76 °C. *R*_f = 0.67 (toluene–ethyl acetate 90 : 10). [α]_D²⁰ = +19.3 (*c* = 0.980; CHCl₃). ¹H NMR (600 MHz, DMSO): 1.41 (9H, s, (CH₃)₃), 3.65 (1H, dd, *J* = 10.5 and 7.8, –CHaHb–Br), 3.75 (1H, dd, *J* = 10.5 and 4.5, –CHaHb–Br), 4.24 (1H, t, *J* ~ 7.0, >CH–), 4.33 (2H, m, –O–CH₂–), 4.34 (1H, ddd, *J* = 8.0, 7.8 and 4.5, >CH–N), 7.89 (1H, d, *J* = 8.0, –NH–CO), 7.33 (2H, m, Ar–H), 7.42 (2H, m, Ar–H), 7.74 (2H, m, Ar–H), 7.89 (2H, m, Ar–H). ¹³C NMR (150.9 MHz, DMSO): 27.73 ((CH₃)₃), 32.81 (–CH₂–Br), 46.77 (>CH–), 56.06 (>CH–N), 66.06 (–CH₂–O–CO), 81.99 (O–C (CH₃)₃), 120.30(2), 125.44(2), 127.25(2) and 127.84(2) (8 \times Ar=CH–), 140.92(2) and 143.90(2) (4 \times Ar>C=), 156.04 (N–CO–O), 168.15 (O–CO–). IR (KBr) ν_{\max} (cm^{–1}) 3349 m (NH); 1741 vs (C=O) ester; 1696 vs C=O (carbamate); 1514 s (amide II); 1247 s, 1159 s (C–O–C); 3066 w, 3041 w, 1478 m, 760 s, 739 s (ring); 2978 m, 1394 m,

1370 m (CH₃). HRMS (ESI) calc for C₂₂H₂₄O₄NBrNa [M + Na]⁺ 468.07809, found: 468.07840.

tert-Butyl 2-(S)-(tert-butoxycarbonylamino)-3-hydroxypropanoate 8

Z-L-Ser-OtBu **4** (12.8 g; 43.3 mmol) was put into a glass pressure bottle and dissolved in 300 ml of methanol, and then 500 mg of 10% Pd/C was added. The mixture was vigorously stirred and allowed to react under the atmosphere of hydrogen (15 psi) at RT overnight. TLC analysis revealed (toluene–ethyl acetate 50 : 50) that the starting compound had completely disappeared. The catalyst was filtered off through celite, and the filter cake was washed in 300 ml of methanol. The filtrate was evaporated *in vacuo* to give 7.2 g of light brown residue, which was immediately dissolved in a 100 ml saturated solution of NaHCO₃. The flask was placed into the ice bath, and di-tert-butyl dicarbonate (Boc₂O) (9.5 g; 43.3 mmol) in 100 ml dioxane was added dropwise under stirring. After the addition of all Boc₂O, stirring was continued for 1 h at 0 °C and then overnight at RT. Thereafter, 500 ml of water was poured in, and the reaction mixture was transferred to the separatory funnel and extracted four times with 150 ml of ethyl acetate. Combined organic layers were washed consecutively twice with 100 ml of water and twice with 100 ml of brine and dried over anhydrous Na₂SO₄. Evaporation of the filtrate under reduced pressure furnished a colorless oil, which was triturated in petroleum ether at –20 °C: colorless solid, yield 9.7 g (86% over two steps). The spectra and physicochemical characteristics of the product are the same as those for **3** prepared by a direct alkylation of Boc-L-Ser.

Alternatively, L-serine **1** (10.5 g; 0.1 mol; [α]_D²⁰ = +13, c = 5, 5 M HCl) was placed into a 1 l, round-bottom flask equipped with a magnetic spin bar. The compound **1** was dissolved in the solution of sodium hydrogen carbonate (16.8 g; 0.2 mol) in 150 ml of water. The flask was immersed into the ice cooling bath, and 1.1 eq. of Boc₂O (24 g; 0.11 mol) in 100 ml of dioxane was added dropwise under vigorous stirring for 30 min. When the addition of Boc anhydride was completed, the reaction mixture was allowed to react for 1 h at 0 °C and then overnight at RT. A saturated aqueous solution of citric acid was added carefully until acidic pH ≈ 3 was reached. The aqueous–organic solution was saturated with sodium chloride, followed by four extractions with ethyl acetate (each with 200 ml). Combined organic phases were washed four times with 100 ml of brine and dried over Na₂SO₄. The filtration of the drying agent, followed by evaporation of the filtrate under reduced pressure, furnished 25 g of the crude [(S)-2-(tert-butoxycarbonylamino)]-3-hydroxypropanoic acid as a colorless oil (R_f = 0.46 ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5), which was used in the next step without additional purification.

Compound **8** was prepared in the same manner as **4** by the reaction of 25 g of Boc-L-Ser, 200 ml of tert-butyl bromide, potassium carbonate (69.1 g; 0.5 mol) and benzyl triethylammonium chloride (CAS 56-37-1, 11.4 g; 0.05 mol) in 450 ml of dimethylacetamide. The required product was a clear oil, which solidified upon standing in a refrigerator at 5 °C. An analytical sample was prepared by trituration in petroleum ether at –20 °C. Yield 18.7 g (72% over two steps), m.p. 77–80 °C. R_f = 0.55 (toluene–ethyl acetate 50 : 50). [α]_D²⁰ = –22.2 (c = 1.928; EtOH). ¹H NMR (600 MHz, DMSO): 1.38 (9H, s, (CH₃)₃), 1.39 (9H, s, (CH₃)₃), 3.60 (2H, dd, J = 6.1 and 5.0, –O–CH₂–), 3.89 (1H, dt, J = 8.2, 5.0 and 5.0, >CH–N), 4.78 (1H, t, J = 6.1, –OH), 6.75 (1H, d, J = 8.2, –NH–CO). ¹³C NMR (150.9 MHz, DMSO): 27.87 ((CH₃)₃), 28.33 ((CH₃)₃), 57.12 (>CH–N); 61.63 (–CH₂–O), 78.32 (O–C(CH₃)₃), 80.53 (O–C(CH₃)₃), 155.49 (N–CO–O), 170.23

(O–CO–). IR (KBr) ν_{max} cm^{–1} 3322 s, 3280 s (NH); 1741 vs (C=O) ester; 1684 vs (C=O) carbamate; 1498 m (amide II); 1155 vs (OC(CH₃)₃); 2976 s, 2933 s, 1395 s, 1366 s (CH₃); 1081 s, 1059 s, 1048 s (C–OH). HRMS (ESI) calc for C₁₂H₂₄O₅N [M + H]⁺ 262.16490, found: 262.16507.

tert-Butyl 2-(R)-(tert-butoxycarbonylamino)-3-hydroxypropanoate 9

With the method described for **8**, the title enantiomer **9** was prepared from D-serine (10.5 g; 0.1 mol; [α]_D²⁰ = –14.75, c = 10; 2 N HCl). Yield 17.5 g (67% over two steps), m.p. 76–78 °C. R_f = 0.55 (toluene–ethyl acetate 50 : 50). [α]_D²⁰ = +21.6 (c = 0.283; EtOH). ¹H NMR (600 MHz, DMSO): 1.38 (9H, s, (CH₃)₃), 1.39 (9H, s, (CH₃)₃), 3.60 (2H, dd, J = 6.0 and 5.0, HO–CH₂–), 3.89 (1H, dt, J = 8.2, 5.0 and 5.0, >CH–N), 4.78 (1H, t, J = 6.0, –OH), 6.76 (1H, d, J = 8.2, –NH–CO). ¹³C NMR (150.9 MHz, DMSO): 27.86 ((CH₃)₃), 28.33 ((CH₃)₃), 57.12 (>CH–N), 61.63 (–CH₂–O), 78.31 (O–C(CH₃)₃), 80.52 (O–C(CH₃)₃), 155.48 (N–CO–O), 170.22 (O–CO–). IR (KBr) ν_{max} (cm^{–1}) 3321 m, 3278 m (NH); 1740 vs (C=O) ester; 1683 vs (C=O) carbamate; 1497 m (amide II); 1155 vs (C–O); 2975 m, 2933 s, 1395 s, 1366 s (CH₃); 1153 vs (OC(CH₃)₃); 1081 s, 1059 s, 1048 s (C–OH). HRMS (ESI) calc for C₁₂H₂₄O₅N [M + Na]⁺ 284.14684, found: 284.14692.

tert-Butyl 2-(tert-butoxycarbonylamino)acrylate **10**, tert-Butyl 2-(S)-(tert-butoxycarbonylamino)-3-azidopropanoate **11** and tert-Butyl 2-(R)-(tert-butoxycarbonylamino)-3-azidopropanoate **12**

Ester **8** (18.7 g; 71.6 mmol) was dissolved in 150 ml of DCM with TEA (10.9 g; 107.4 mmol). The reaction mixture was cooled in the ice bath, and MsCl (9 g; 78.8 mmol) was added dropwise. After half an hour of cooling and stirring, TLC analysis revealed that starting compound **8** had been consumed. Except for mesyl intermediate (R_f = 0.73 in toluene–ethyl acetate, 80 : 20), the elimination product **10** (R_f = 0.88 in toluene–ethyl acetate, 80 : 20) was also detected on the TLC plate. Excess of the base was eliminated by the addition of 1 M solution of citric acid; the lower organic layer was washed with 50 ml of water and 50 ml of brine and dried over Na₂SO₄. Volatile materials were evaporated under reduced pressure to give 23.8 g of brown oil. The crude intermediate was dissolved in 100 ml DMSO with NaN₃ (9.1 g; 140.4 mmol) and heated at 70 °C overnight. After cooling, 200 ml of water was added, and the solution was extracted four times with 200 ml of diethyl ether. The combined organic layers were washed twice with 100 ml of water and twice with 100 ml brine and dried over Na₂SO₄. The drying agent was filtered off, and the filtrate was evaporated under reduced pressure to afford 20 g of yellow residue, which was purified by flash chromatography on silica gel, using a linear gradient of diethyl ether in petroleum ether. Two major compounds were isolated, the product of elimination **10** and the required azide **11**. Yield **10** 2.7 g (16% over two steps). Yield **11** 9.7 g (47% over two steps).

Alternatively, ester **8** (7.3 g; 27.9 mmol) was treated in 50 ml DCM with (3.3 g; 41.9 mmol) pyridine and (3.5 g; 30.7 mmol) MsCl, followed by the addition of NaN₃ (3.5 g; 53.1 mmol) in 50 ml DMSO by the protocol described earlier. The starting compound was fully reacted after 24 h; the formation of **10** during the mesylation step was also observed. Yield **10** 1.8 g (26% over two steps), yield **11** 3.7 g (46% over two steps).

Alternatively, **13** (9.7 g; 29.9 mmol) was dissolved in 50 ml of DMSO with NaN₃ (3.9 g; 59.8 mmol) and heated at 70 °C overnight. After cooling, 100 ml of water was added, and the solution was extracted four times with 50 ml of diethyl ether. The combined organic layers were washed twice with 50 ml of water and twice with 50 ml of brine and dried over Na₂SO₄. The drying agent was

filtered off, and the filtrate was evaporated under reduced pressure to afford 8.6 g of dark yellow residue, which was purified by flash chromatography on silica gel, using a linear gradient of diethyl ether in petroleum ether. Yield **10** 2.9 g (40%), yield **11** 3.5 g (41%).

Using the method described for **11**, the title enantiomer **12** was prepared from **9** (10.5 g; 40.2 mmol), MsCl (5.1 g; 44.2 mmol), TEA (6.1 g; 60.3 mmol) and NaN₃ (5.3 g; 80.4 mmol). Yield **10** 2.1 g (21%), yield **12** 5.8 g (50%).

Compound **10**. Colorless oil. $R_f = 0.77$ (toluene–ethyl acetate 90 : 10). ¹H NMR (600 MHz, CDCl₃): 1.47 (9H, s, (CH₃)₃), 1.51 (9H, s, (CH₃)₃), 5.63 and 6.07 (2H, 2 × d, $J = 1.5$, =CH₂), 7.03 (1H, br, s, –NH–CO). ¹³C NMR (150.9 MHz, CDCl₃): 27.88 ((CH₃)₃), 28.22 ((CH₃)₃), 80.39 (O–C(CH₃)₃), 82.54 (O–C(CH₃)₃), 103.95 (=CH₂), 132.43 (=C<), 152.60 (N–CO–O), 163.02 (O–CO–). IR (CCl₄) ν_{\max} (cm⁻¹) 3461 w, 3421 m (NH); 1736 vs (C=O) ester; 1707 vs (C=O) carbamate; 1509 vs (amide II); 1157 vs, 1065 vs (C–O–C); 2980 s, 2934 s, 1393 s, 1369 s (CH₃); 3005 w, 1641 m (C=C). HRMS (ESI) calc for C₁₂H₂₁O₄NNa [M + Na]⁺ 266.13628, found: 266.13635.

Compound **11**. Colorless liquid. $R_f = 0.59$ (toluene–ethyl acetate 90 : 10). $[\alpha]_D^{20} = +26.9$ ($c = 0.304$; CHCl₃). ¹H NMR (600 MHz, CDCl₃): 1.36 (9H, s, (CH₃)₃), 1.40 (9H, s, (CH₃)₃), 3.57 (1H, dd, $J = 12.4$ and 3.7, –CHaHb–N₃), 3.62 (1H, dd, $J = 12.4$ and 3.5, –CHaHb–N₃), 4.25 (1H, ddd, $J = 7.4$, 3.7 and 3.5, >CH–N), 5.26 (1H, d, $J = 7.4$, –NH–CO). ¹³C NMR (150.9 MHz, CDCl₃): 27.92 ((CH₃)₃), 28.27 ((CH₃)₃), 53.00 (–CH₂–N₃), 54.13 (>CH–N); 80.19 (O–C(CH₃)₃), 83.13 (O–C(CH₃)₃), 155.08 (N–CO–O), 168.72 (O–CO–). IR (CCl₄) ν_{\max} (cm⁻¹) 3435 m, 3364 m (NH); 2108 vs (N₃); 1740 vs (C=O) ester; 1719 vs (C=O) carbamate; 1498 s (amide II); 1155 vs (C(OCH₃)₃); 2980 s, 2935 s, 1394 s, 1369 s (CH₃). HRMS (ESI) calc for C₁₂H₂₂O₄N₄Na [M + Na]⁺ 309.15333, found: 309.15336.

Compound **12**. Colorless liquid. $R_f = 0.59$ (toluene–ethyl acetate 90 : 10). $[\alpha]_D^{20} = -27.9$ ($c = 0.369$; CHCl₃). ¹H NMR (600 MHz, CDCl₃): 1.44 (9H, s, (CH₃)₃), 1.48 (9H, s, (CH₃)₃), 3.65 (1H, dd, $J = 12.3$ and 3.6, –CHaHb–N₃), 3.70 (1H, dd, $J = 12.3$ and 3.6, –CHaHb–N₃), 4.33 (1H, dt, $J = 7.3$, 3.6 and 3.6, >CH–N), 5.35 (1H, br d, $J = 7.3$, –NH–CO). ¹³C NMR (150.9 MHz, CDCl₃): 27.91 ((CH₃)₃), 28.26 ((CH₃)₃), 52.99 (–CH₂–N₃), 54.13 (>CH–N), 80.17 (O–C(CH₃)₃), 83.11 (O–C(CH₃)₃), 155.08 (N–CO–O), 168.71 (O–CO–). IR (CCl₄) ν_{\max} (cm⁻¹) 3435 m (NH); 2109 vs (N₃); 1741 vs (C=O) ester; 1716 vs (C=O) carbamate; 1492 s (amide II); 1155 vs (C(OCH₃)₃); 2981 s, 2933 s, 1393 s, 1369 s (CH₃). HRMS (ESI) calc for C₁₂H₂₂O₄N₄Na [M + Na]⁺ 309.15333, found: 309.15342.

tert-Butyl 2-(S)-(tert-butoxycarbonylamino)-3-bromopropanoate **13**

Ester **8** (11.5 g; 44 mmol) and CBr₄ (17.5 g; 52.8 mmol) were dissolved in 100 ml of DCM. The flask with the reaction mixture was immersed in the ice cooling bath, and PPh₃ (13.8 g; 52.8 mmol) in 50 ml of DCM was added dropwise under stirring. Stirring was continued for 1 h at 0 °C and then at RT overnight. DCM was removed under reduced pressure, and the brown residue was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in petroleum ether. The product was a colorless oil, which was triturated in petroleum ether at –20 °C. Yield 9.7 g (68%). White solid, m.p. 64–65 °C. $R_f = 0.65$ (toluene–ethyl acetate 90 : 10). $[\alpha]_D^{20} = -7$ ($c = 0.284$; CHCl₃). ¹H NMR (600 MHz, DMSO): 1.39 (9H, s, (CH₃)₃), 1.41 (9H, s, (CH₃)₃), 3.61 (1H, dd, $J = 10.4$ and 7.8, –CHaHb–Br), 3.70 (1H, dd, $J = 10.4$ and 4.5, –CHaHb–Br), 4.22 (1H, ddd, $J = 8.2$, 7.8 and 4.5, >CH–N), 7.21 (1H, d, $J = 8.2$, –NH–CO). ¹³C NMR (150.9 MHz, DMSO): 27.72 ((CH₃)₃), 28.27 ((CH₃)₃),

32.89 (–CH₂–Br), 55.75 (>CH–N); 78.77 (O–C(CH₃)₃), 81.75 (O–C(CH₃)₃), 155.29 (N–CO–O), 168.41 (O–CO–). IR (KBr) ν_{\max} (cm⁻¹) 3433 m (NH); 1736 vs (C=O) ester; 1714 vs (C=O) carbamate; 1498 s (amide II); 1157 s, 1067 s (C–O–C); 2980 s, 2934 s, 1395 s, 1369 s (CH₃). HRMS (ESI) calc for C₁₂H₂₂O₄NBrNa [M + Na]⁺ 346.06244, found: 346.06254.

2-(S)-(9-Fluorenylmethoxycarbonylamino)-3-azidopropanoic acid (Fmoc-β-azido-L-Ala) **14** Compound **11** (9.6 g; 33.5 mmol) was treated with the cleavage cocktail, consisting of 18.8 ml of DCM, 18.8 ml of TFA and 2.4 ml of water. The reaction started with severe liberation of CO₂ and isobutene and continued at RT overnight under stirring. Volatile materials were removed on the rotary evaporator. The yellow residue was dissolved in a solution of NaHCO₃ (11.3 g; 134 mmol) in 50 ml of water. The reaction mixture was cooled in the ice bath, and Fmoc-OSu (11.3 g; 33.5 mmol) in 50 ml dioxane was added dropwise under vigorous stirring. The reaction mixture was allowed to react for 1 h at 0 °C and then overnight at RT. The flask was again ice cooled, and concentrated hydrochloric acid was carefully added until acidic pH ≈ 1 was reached. The reaction mixture was extracted thrice with 100 ml of ethyl acetate. Thereafter, the combined organic layers were successively washed once with 100 ml of water and twice with 100 ml of brine, followed by drying on Na₂SO₄. Evaporation of the filtrate gave a brown oil, which was purified by flash chromatography on silica gel, using a linear gradient of 1% CH₃COOH/ethyl acetate in toluene. Evaporation of the product afforded a yellow semisolid, which was triturated in toluene at –20 °C. Yield 10.2 g (86%). Colorless solid. m.p. 119–120 °C. $R_f = 0.63$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = -4.8$ ($c = 0.271$; DMF). ¹H NMR (600 MHz, DMSO): 3.53 (1H, dd, $J = 12.1$ and 6.4, –CHaHb–N₃), 3.68 (1H, dd, $J = 12.1$ and 3.6, –CHaHb–N₃), 3.98 (1H, ddd, $J = 7.0$, 6.4 and 3.6, >CH–N), 4.22 (1H, m, >CH–), 4.22 and 4.33 (2H, 2 × m, –O–CH₂–), 6.95 (1H, d, $J = 7.0$, –NH–CO), 7.31 (2H, m, Ar–H), 7.40 (2H, m, Ar–H), 7.69 (2H, m, Ar–H), 7.87 (2H, m, Ar–H). ¹³C NMR (150.9 MHz, DMSO): 46.92 (>CH–), 52.89 (–CH₂–N₃), 56.08 (>CH–N), 65.86 (–CH₂–O–CO), 120.30(2), 125.40, 125.46, 127.29(2) and 127.82 (2) (8 × Ar=CH–), 140.90(2), 144.01 and 144.10 (4 × Ar>C=), 155.83 (N–CO–O), 172.29 (–COOH). IR (KBr) ν_{\max} (cm⁻¹) 3402 m (NH); 2108 vs (N₃); 1719 vs (C=O) acid; 1693 vs (C=O) carbamate; 1539 s (amide II); 3065 w, 3041 w, 1603 vs, 1478 s, 1451 s, 757 m, 740 s (ring). HRMS (ESI) calc for C₁₈H₁₃O₄N₄[M – H]⁺ 351.10988, found: 351.10968.

Alternatively, compound **14** was prepared from **27**. Triflic anhydride (13.3 g; 47.2 mmol) was added dropwise under ice cooling and vigorous stirring to the two-phase system of NaN₃ (15.3 g; 236 mmol) in 60 ml of water and 70 ml of DCM. The ice bath was removed and stirring continued for 2 h. The aqueous layer was separated and extracted twice with 50 ml of DCM. Thereafter, the combined organic phases were washed with 5% NaHCO₃. The resulting solution of triflic azide (TFN₃) in DCM was immediately added dropwise to the suspension of **27** (7.7 g; 23.6 mmol), NaHCO₃ (19.8 g; 236 mmol) and CuSO₄·5H₂O (60 mg; 23.6 μmol) in 50 ml of water and 150 ml of methanol, and the mixture was stirred at RT overnight. Volatile material was evaporated, and the remaining slurry was carefully acidified with concentrated HCl until pH 1–2 was reached. The reaction mixture was extracted four times with 100 ml of ethyl acetate. The combined organic layers were washed twice with 100 ml of water and twice with 100 ml of brine and dried over Na₂SO₄. Filtering off the drying agent and evaporation of the filtrate gave 8 g of brown residue, which was purified by flash chromatography on silica gel, using a linear gradient of 1% CH₃COOH/ethyl acetate in toluene. The yellow oil was triturated

in toluene at $-20\text{ }^{\circ}\text{C}$ to give the pure product. Yield 6.6 g (80%). Physicochemical characteristics were consistent with the above-listed ones.

Alternatively, compound **14** was prepared from **32**. Compound **32** (8.4 g; 36.5 mmol) was treated with a mixture of 18.8 ml of DCM, 18.8 ml of TFA and 2.4 ml of water. After 2 h of stirring, volatile materials were evaporated, and the yellow oil was dissolved in 50 ml of water with NaHCO_3 (9.2 g; 109.5 mmol). The flask was immersed in an ice cooling bath, and Fmoc-OSu (12.3 g; 36.5 mol) in 100 ml of dioxane was added dropwise during a period 15 min under vigorous stirring. When the addition of Fmoc-OSu was complete, the slurry was allowed to react for 1 h at $0\text{ }^{\circ}\text{C}$ and then overnight at RT. The reaction mixture was again cooled in an ice bath, and concentrated HCl was added dropwise until pH ~ 0 –1 was reached. Thereafter, 150 ml of water was added, and the reaction mixture was extracted thrice with 100 ml of ethyl acetate. The combined organic layers were washed once with 100 ml of water and twice with 100 ml of brine and dried over Na_2SO_4 . The filtrate was evaporated, and the resulting brown oil was subjected to flash chromatography on silica gel, using a linear gradient of 1% CH_3COOH /ethyl acetate in toluene. The yellow oil was triturated in toluene at $-20\text{ }^{\circ}\text{C}$ to afford the pure product. Yield 8.2 g (64%). Physicochemical characteristics were consistent with the above-listed ones.

2-(R)-(9-Fluorenylmethyloxycarbonylamino)-3-azidopropanoic Acid (Fmoc- β -azido-D-Ala) 15

With the method described for **14**, the title enantiomer **15** was prepared from **12** (7.4 g; 25.9 mmol), 18.8 ml TFA, NaHCO_3 (8.7 g; 103.6 mmol) and Fmoc-OSu (8.7 g; 25.9 mmol). Yield 7.8 g (86%). Colorless solid. m.p. 117 – $119\text{ }^{\circ}\text{C}$. $R_f = 0.63$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = +5.1$ ($c = 0.235$; DMF). ^1H NMR (600 MHz, DMSO): 3.61 (1H, dd, $J = 13.0$ and 5.2 , $-\text{CHaHb-N}_3$), 3.64 (1H, dd, $J = 13.0$ and 7.1 , $-\text{CHaHb-N}_3$), 4.24 (1H, ddd, $J = 8.3$, 7.1 and 5.2 , $>\text{CH-N}$), 4.24 (1H, m, $>\text{CH-}$), 4.31 (1H, dd, $J = 10.5$ and 6.7 , $-\text{CHaHb-O}$), 4.33 (1H, dd, $J = 10.5$ and 7.5 , $-\text{CHaHb-O}$), 7.33 (2H, m, Ar-H), 7.42 (2H, m, Ar-H), 7.74 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 7.92 (1H, d, $J = 8.3$, $-\text{NH-CO}$). ^{13}C NMR (150.9 MHz, DMSO): 46.79 ($>\text{CH-}$), 51.12 ($-\text{CH}_2-\text{N}_3$), 54.01 ($>\text{CH-N}$), 66.03 ($-\text{CH}_2-\text{O-CO}$), 120.29(2), 125.45(2), 127.26(2) and 127.84(2) ($8\times \text{Ar}=\text{CH-}$), 140.90(2), 143.92 and 143.96 ($4\times \text{Ar}>\text{C=}$), 156.22 (N-CO-), 171.23 (COOH). IR (KBr) ν_{max} (cm^{-1}) 3403 m (NH); 2109 vs (N_3); 1712 vs (C=O) acid; 1707 vs (C=O) carbamate; 1538 s (amide II); 3065 w, 3041 w, 1611 w, 1580 w, 1478 m, 1451 s, 758 m, 740 s (ring). HRMS (ESI) calc for $\text{C}_{18}\text{H}_{16}\text{O}_4\text{N}_4\text{Na}[\text{M} + \text{Na}]^+$ 375.10638, found: 375.10653.

2-(S)-Amino-3-azidopropanoic Acid Trifluoroacetic Acid Salt 16

One gram of 2-Cl-Trt-chloride resin (Merck Novabiochem, Darmstadt, Germany, capacity 1.5 mmol/g, 100–200 mesh) was placed in a 20 ml syringe with a frit and preswollen in 10 ml DMF for half an hour. The solvent was removed, and **14** (0.575 g; 1.5 mmol) in 4 ml of DMF and DIPEA (783 μl ; 4.5 mmol) in 2 ml of DMF were added. The syringe was agitated by shaking for 1.5 h, followed by washing (3×10 ml of DMF). The reaction was terminated by two subsequent additions of a mixture of 5.1 ml of DCM, 0.6 ml of CH_3OH and 0.3 ml of DIPEA, each for 5 min. The resin was washed thrice with 10 ml of DCM and thrice with 10 ml of DMF. The Fmoc group was cleaved with 20% (v/v) piperidine in DMF (10 ml for 5 and 30 min). The resin was washed thrice with 10 ml of DMF and

thrice with 10 ml of DCM. Finally, the product was cleaved from the resin by three subsequent treatments with a mixture of 2 ml of AcOH, 2 ml of trifluoroethanol, CAS 75-89-8, and 6 ml of DCM, each for 15 min. Filtrates were evaporated to dryness, and the crude material was subjected to RP-HPLC. The following gradient was used: $t = 0$ min (2% B), $t = 15$ min (15% B), $t = 31$ min (100% B). Yield 270 mg (74%). White lyophilisate. $[\alpha]_D^{20} = +21.8$ ($c = 0.262$; H_2O). ^1H NMR (600 MHz, $\text{D}_2\text{O} + \text{NaOD}$): 3.51 (1H, dd, $J = 5.6$ and 4.3 , $>\text{CH-N}$), 3.60 (1H, dd, $J = 12.6$ and 4.3 , $-\text{CHaHb-N}_3$), 3.65 (1H, dd, $J = 12.6$ and 5.6 , $-\text{CHaHb-N}_3$). IR (KBr) ν_{max} (cm^{-1}) 2122 vs (N_3); 1733 s (C=O) acid; 1640 s, 1443 m (C=O) CF_3COO^- ; 1619 m, 1535 m (NH_3^+); 1207 s, 1151 m (CF). HRMS (ESI) calc for $\text{C}_3\text{H}_7\text{O}_2\text{N}_4[\text{M} + \text{H}]^+$ 131.05635, found: 131.05640.

Fmoc- β -azido-Ala-Val-OBn 18

Fmoc- β -azido-Ala **14** (326 mg; 1 mmol), L-Val-OBn-TsOH (379 mg; 1 mmol), HOBt-H $_2\text{O}$ (135 mg; 1 mmol), bromotripyrrolidinophosphonium hexafluorophosphate (CAS 132705-51-2, 606 mg; 1.3 mmol) and DIPEA (387 mg; 3 mmol) were stirred at RT in 7 ml of DMF. The solvent was evaporated under reduced pressure, and the residue was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in toluene. Yield 400 mg (74%). Semisolid. $R_f = 0.54$ (toluene–ethyl acetate 80 : 20). $[\alpha]_D^{20} = -7.4$ ($c = 0.270$; DMF). ^1H NMR (600 MHz, DMSO): 0.862 (3H, d, $J = 6.8$, $-\text{CH}_3$), 0.866 (3H, d, $J = 6.8$, $-\text{CH}_3$), 2.08 (1H, m, $-\text{CH-}$), 3.39 (1H, dd, $J = 12.5$ and 9.1 , $-\text{CHaHb-N}_3$), 3.47 (1H, dd, $J = 12.5$ and 4.4 , $-\text{CHaHb-N}_3$), 4.22–4.31 (3H, m, $>\text{CH-CH}_2-\text{O-CO}$), 4.24 (1H, t, $J = 8.2$, $>\text{CH-N}$), 4.43 (1H, ddd, $J = 9.1$, 8.7 and 4.4 , $>\text{CH-N}$), 5.10 and 5.15 (2H, 2 \times d, $J = 12.4$, $-\text{O-CH}_2-$), 7.32 (2H, m, Ar-H), 7.33–7.37 (5H, m, $-\text{C}_6\text{H}_5$), 7.42 (2H, m, Ar-H), 7.73 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 7.81 (1H, d, $J = 8.7$, $-\text{NH-CO}$), 8.36 (1H, d, $J = 8.2$, $-\text{NH-CO}$). ^{13}C NMR (150.9 MHz, DMSO): 18.21 (CH_3), 19.06 (CH_3), 27.81 ($(\text{CH}_3)_3$), 30.02 ($>\text{CH-}$), 46.78 ($>\text{CH-}$), 51.69 ($-\text{CH}_2-\text{N}_3$), 54.43 ($>\text{CH-N}$), 57.79 ($>\text{CH-N}$), 66.04 ($-\text{CH}_2-\text{O-CO}$), 66.22 ($-\text{CH}_2-\text{O-CO}$), 120.29(2), 125.48, 125.51, 127.25(2) and 127.84(2) ($8\times \text{Ar}=\text{CH-}$), 128.20–128.60(5) ($-\text{C}_6\text{H}_5$), 140.90(2), 143.91 and 143.97 ($4\times \text{Ar}>\text{C=}$), 156.10 (N-CO-), 169.65 (N-CO-), 171.17 (N-CO-). IR (KBr) ν_{max} (cm^{-1}) 3291 vs (NH); 2105 vs (N_3); 1733 vs (C=O) ester; 1690 vs (C=O) carbamate; 1655 vs (C=O) amide; 1539 vs (amide II); 1252 s (C–O–C) ester; 3066 w, 3038 w, 1603 vs, 1478 m 1464 m, 1450 s, 758 s, 740 s (ring). HRMS (ESI) calc for $\text{C}_{30}\text{H}_{31}\text{O}_5\text{N}_5\text{Na}[\text{M} + \text{Na}]^+$ 564.22174, found: 564.22118.

Experimental procedures and analytical data for compound **20** are provided in the supporting information.

Fmoc- β -azido-Ala-Val-OtBu 21

Fmoc- β -azido-Ala **14** (0.652 g; 2 mmol), L-Val-OtBu-TsOH **20** (0.419 g; 2 mmol), bromotripyrrolidinophosphonium hexafluorophosphate (1.17 g; 2.5 mmol) and DIPEA (1 g; 8 mmol) in 10 ml of DMF were stirred overnight at RT. The solvent was evaporated under reduced pressure, and the residue was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in toluene. An analytical sample was gained by crystallization from a mixture of ethyl acetate–petroleum ether. Yield 700 mg (70%). $R_f = 0.82$ (toluene–ethyl acetate 50 : 50). White solid, 154 – $155\text{ }^{\circ}\text{C}$. $[\alpha]_D^{20} = -4.2$ ($c = 0.357$; DMF). ^1H NMR (600 MHz, DMSO): 0.87 (3H, d, $J = 6.9$, CH_3), 0.88 (3H, d, $J = 6.9$, CH_3), 1.39 (9H, s, $(\text{CH}_3)_3$), 2.04 (1H, m, $>\text{CH-}$), 1.41 (9H, s, $(\text{CH}_3)_3$), 3.44 (1H, dd, $J = 12.5$ and 8.8 , $-\text{CHaHb-N}_3$), 3.54 (1H, dd, $J = 12.5$ and 4.2 , $-\text{CHaHb-N}_3$), 4.23 (1H, dd, $J = 7.5$ and 6.9 , $>\text{CH-}$), 4.27 (1H, dd, $J = 10.4$ and 6.9 , $-\text{CHaHb-N}_3$).

CHaHb-O-CO), 4.31 (1H, dd, $J = 10.4$ and 7.5 , -CHaHb-O-CO), 4.44 (1H, td, $J = 8.8$ and 4.2 , >CH-N), 7.84 (1H, d, $J = 8.8$, -NH-CO), 8.20 (1H, d, $J = 8.2$, -NH-CO), 7.32 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.73 (2H, m, Ar-H), 7.89 (2H, m, Ar-H). ^{13}C NMR (150.9 MHz, DMSO): 18.90 (CH_3), 19.10 (CH_3), 27.81 ($(\text{CH}_3)_3$), 30.13 (>CH-), 46.79 (>CH-), 51.81 (-CH₂-N₃), 54.49 (>CH-N), 58.19 (>CH-N), 66.08 (-CH₂-O-CO), 80.95 (O-C(CH₃)₃), 120.33(2), 125.52(2), 127.28(2) and 127.87(2) (8 \times Ar=CH-), 140.92(2), 143.94 and 143.99 (4 \times Ar>C=), 156.14 (N-CO-O), 169.45 (N-CO-), 170.43 (-O-CO-). IR (KBr) ν_{max} (cm^{-1}) 3335 s, 3266 s (NH); 2964 s, 2873 w, 1479 m, 1386 s, 1370 s (CH_3); 2106 vs (N_3); 1730 vs (C=O) ester; 1714 vs (C=O) carbamate; 1659 vs (C=O) amide; 1555 s, 1533 vs (amide II); 1253 vs (C-O); 1160 s (-C(CH₃)₃); 1143 s (-CH(CH₃)₂); 3069 w, 3045 w, 3007 w, 1479 m, 1451 s, 1312 s, 1110 m, 799 m, 761 s, 742 vs (ring). HRMS (ESI) calc for $\text{C}_{22}\text{H}_{33}\text{O}_3\text{N}_3\text{Na}[\text{M} + \text{Na}]^+$ 530.23739, found: 530.23732.

β -Azido-Ala-Val 22

Compound **21** (600 mg; 1.2 mmol) was treated with a mixture consisting of 2 ml of DCM, 2.5 ml of TFA and 40 μl of water. After 2 h, TLC analysis revealed completely deprotected *tert*-butyl moiety, and the volatile material was evaporated under reduced pressure. 2-Cl-Trt-chloride resin (0.8 g, Merck Novabiochem, capacity 1.5 mmol/g, 100–200 mesh) was placed into a 20 ml syringe with a frit and preswollen in 10 ml of DMF for half an hour. DMF was removed, and the crude acid Fmoc- β -azido-Ala-Val in 4 ml DMF and DIPEA (627 μl ; 3.6 mmol) in 2 ml DMF were added. The syringe was agitated by shaking for 1.5 h, followed by washing thrice with 10 ml of DMF. In the next step, the Fmoc group was cleaved with 20% (v/v) piperidine in DMF (10 ml for 5 and 30 min). The resin was washed thrice with 10 ml of DMF and thrice with 10 ml of DCM. Finally, the product was cleaved from resin by three subsequent treatments with 10 ml of AcOH, 15 min each. Filtrates were combined and evaporated to dryness. The residue was sonicated for 10 min in 15 ml of diethyl ether. The white precipitate was decanted and dissolved in 2 ml of hot acetonitrile (60 °C). After cooling, the crystals were filtered off and washed with 5 ml of diethyl ether. Yield 181 mg (67%). White solid, m.p. 187–189 °C. $[\alpha]_{\text{D}}^{20} = +8.9$ ($c = 0.357$; DMSO). ^1H NMR (600 MHz, DMSO): 0.856 (3H, d, $J = 6.8$, CH_3), 0.865 (3H, d, $J = 6.8$, CH_3), 2.07 (1H, m, >CH-), 3.44 (1H, dd, $J = 12.4$ and 6.4 , -CHaHb-N₃), 3.48 (1H, dd, $J = 12.4$ and 4.4 , -CHaHb-N₃), 3.55 (1H, dd, $J = 6.4$ and 4.4 , >CH-N), 4.13 (1H, dd, $J = 8.6$ and 5.2 , >CH-N), 8.17 (1H, d, $J = 8.6$, -NH-CO) ^{13}C NMR (150.9 MHz, DMSO): 18.01 (CH_3), 19.43 (CH_3), 30.59 (>CH-), 54.34 (>CH-N), 54.49 (-CH₂-N₃), 57.46 (>CH-N), 171.53 (N-CO-), 173.16 (-COOH). IR (KBr) ν_{max} (cm^{-1}) 3433 m + vbr (NH); 2967 s, 2875 m, 1389 s, 1376 m (CH_3); 2938 m, 1443 m (CH_2); 2110 vs (N_3); 1671 vs (C=O) amide; 1506 vs (amide II); 1581 s; 1403 m (COO⁻); 1174 m (-CH(CH₃)₂). HRMS (ESI) calc for $\text{C}_8\text{H}_{14}\text{O}_3\text{N}_3[\text{M} - \text{H}]^-$ 228.11021, found: 228.10999.

Experimental procedures and analytical data for compounds **25** and **26** are provided in the supporting information.

2-(5)-(9-Fluorenylmethyloxycarbonylamino)-3-aminopropionic Acid 27

Compound **25** (11.1 g; 31.3 mmol) was suspended in a mixture of 48 ml of acetonitrile, 48 ml of ethyl acetate and 24 ml of water. (Diacetoxiodo)benzene (12.1 g; 37.6 mmol) was added under vigorous stirring in five portions, and the resulting slurry was allowed to react overnight. The flask was ice cooled, and the crystals were filtered off in a Büchner funnel and washed with 100 ml of chilled

ethyl acetate. Final purification was carried out by dissolving the solid in 150 ml of hot isopropyl alcohol (70 °C). After cooling, the crystals were filtered off and dried under deep vacuum. Yield 7.7 g (75%). White solid, m.p. 130–132 °C. $R_f = 0.75$ (ethyl acetate–acetone–methanol–water 4 : 1 : 1 : 1). $[\alpha]_{\text{D}}^{20} = +4.6$ ($c = 0.306$; acetic acid). ^1H NMR (600 MHz, DMSO): 2.87 and 3.04 (2H, 2 \times m, -CH₂-N), 3.80 (1H, m, >CH-N), 4.20–4.28 (3H, m, >CH-CH₂-O-CO), 6.91 (1H, br s, -NH-CO), 7.32 (2H, m, Ar-H), 7.40 (2H, m, Ar-H), 7.70 (2H, m, Ar-H), 7.88 (2H, m, Ar-H). ^{13}C NMR (150.9 MHz, DMSO): 40.89 (-CH₂-NH₂), 46.84 (>CH-), 52.00 (>CH-N), 65.92 (-CH₂-O-CO), 120.28(2), 125.44, 125.47, 127.29(2) and 127.81(2) (8 \times Ar=CH-), 140.89(2), 144.04 and 144.10 (4 \times Ar>C=), 155.91 (N-CO-O), 171.56 (-COOH). IR (KBr) ν_{max} (cm^{-1}) 3409 s (NH); 1703 vs + vbr (C=O) acid and carbamate; 1530 m (amide II); 3065 w, 3040 w, 1479 m, 1450 m, 759 m, 740 s (ring). HRMS (ESI) calc for $\text{C}_{18}\text{H}_{18}\text{O}_4\text{N}_2\text{Na}[\text{M} + \text{Na}]^+$ 349.11588, found: 349.11595.

2-(5)-(9-Fluorenylmethyloxycarbonylamino)-4-aminobutanoic Acid 28

Compound **28** was prepared by the reaction of **26** (6.6 g; 17.9 mmol) and PIDA (6.9 g; 21.5 mmol) by the method used previously for **27**. Yield 3.4 g (56%). $R_f = 0.70$ (ethyl acetate–acetone–methanol–water 4 : 1 : 1 : 1). $[\alpha]_{\text{D}}^{20} = +22.4$ ($c = 0.241$; DMSO). ^1H NMR (500 MHz, DMSO): 1.87 (2H, m, -CH₂-), 2.87 (2H, t, $J = 6.5$, -CH₂-N), 3.71 (1H, q, $J = 6.3$, >CH-N), 4.20 (1H, m, >CH-), 4.24 (2H, m, -CH₂-O), 6.76 (1H, d, $J = 6.3$, -NH-CO), 7.32 (2H, m, Ar-H), 7.40 (2H, m, Ar-H), 7.68 (2H, m, Ar-H), 7.87 (2H, m, Ar-H). ^{13}C NMR (125.7 MHz, DMSO): 30.97 (-CH₂-), 37.32 (-CH₂-N), 46.97 (>CH-), 54.70 (>CH-N), 65.84 (-CH₂-O-CO), 120.40(2), 125.54(2), 127.42(2) and 127.93(2) (8 \times Ar=CH-), 140.99(2) and 144.25(2) (4 \times Ar>C=), 155.70 (N-CO-O), 173.64 (-COOH). IR (KBr) ν_{max} (cm^{-1}) 3384 m (NH); 1727 vs (C=O) acid; 1705 vs (C=O) carbamate; 1518 s (amide II); 3039 w, 3020 w, 1593 s, 1465 m, 1450 m, 1107 m, 1033 m, 760 m, 736 s (ring); 1478 m (CH_2). HRMS (ESI) calc for $\text{C}_{19}\text{H}_{21}\text{O}_4\text{N}_2\text{Na}[\text{M} + \text{Na}]^+$ 363.13153, found: 363.13162.

2-(5)-(9-Fluorenylmethyloxycarbonylamino)-4-azidobutanoic Acid 29

With the protocol previously employed for **11**, the reaction of **28** (4.5 g; 13.2 mmol), NaHCO_3 (11.1 g; 132 mmol) and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (32 mg; 132 μmol) in a mixture of 75 ml of methanol and 25 ml of water with TfN_3 in dichloromethane gave azide **29**. TfN_3 was prepared by the reaction of NaN_3 (8.6 g; 132 mmol) and triflic anhydride (7.5 g; 26.4 mmol) in 40 ml of water and 60 ml of DCM. Yield 4.4 g (92%). $R_f = 0.68$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_{\text{D}}^{20} = -17.3$ ($c = 0.294$; CH_3OH). ^1H NMR (500 MHz, DMSO): 1.84 and 1.97 (2H, 2 \times m, -CH₂-), 3.34 (1H, ddd, $J = 12.5$, 8.2 and 6.5, -CHaHb-N₃), 3.44 (1H, dd, $J = 12.5$, 7.1 and 5.3, -CHaHb-N₃), 4.04 (1H, ddd, $J = 10.0$, 8.2 and 4.4, >CH-N), 4.23 (1H, m, >CH-), 4.32 (2H, m, -CH₂-O), 7.33 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.70 (1H, d, $J = 8.2$, -NH-CO), 7.71 (2H, m, Ar-H), 7.88 (2H, m, Ar-H). ^{13}C NMR (125.7 MHz, DMSO): 30.20 (-CH₂-), 46.94 (>CH-), 47.85 (-CH₂-N₃), 51.50 (>CH-N), 65.91 (-CH₂-O-CO), 120.42, 120.44, 125.51, 125.53, 127.39(2) and 127.97(2) (8 \times Ar=CH-), 141.02(2), 144.04 and 144.08 (4 \times Ar>C=), 156.47 (N-CO-O), 173.67 (-COOH). IR (KBr) ν_{max} (cm^{-1}) 3333 m (NH); 2950 w, 1478 w (CH_2); 2107 vs (N_3); 1717 vs (C=O) acid; 1697 vs (C=O) carbamate; 1540 s (amide II); 3065 w, 3040 w, 3019 w, 1451 m, 758 m, 739 s (ring). HRMS (ESI) calc for $\text{C}_{19}\text{H}_{18}\text{O}_4\text{N}_4\text{Na}[\text{M} + \text{Na}]^+$ 389.12203, found: 389.12226.

Experimental procedures and analytical data for compound **30** are provided in the supporting information.

2-(5)-(tert-Butoxycarbonylamino)-3-aminopropionic Acid **31**

Intermediate **30** (18.9 g; 81.4 mmol) was suspended in a mixture of acetonitrile (90 ml), 90 ml ethyl acetate (90 ml) and water (45 ml), and PIDA (31.4 g; 97.7 mmol) was added in five portions during 15 min. Then, 10 min after the addition of all amount of PIDA, the slurry turned clear followed by a rapid precipitation of the crude product. The cake of the filtrate was washed out with 200 ml of chilled ethyl acetate, and no additional purification was needed. Yield 12.7 g (77%). White solid, m.p. 209–211 °C. $R_f = 0.40$ (ethyl acetate–acetone–methanol–water 4 : 1 : 1 : 1). $[\alpha]_D^{20} = -5.3$ ($c = 0.318$; acetic acid). $^1\text{H NMR}$ (600 MHz, DMSO): 1.39 (9H, s, $(\text{CH}_3)_3$), 2.77 (1H, dd, $J = 11.9$ and 9.1, $-\text{CHaHb}-\text{N}$), 3.01 (1H, dd, $J = 11.9$ and 5.4, $-\text{CHaHb}-\text{N}$), 3.67 (1H, ddd, $J = 9.1$, 6.0 and 5.4, $>\text{CH}-\text{N}$), 6.28 (1H, br d, $J = 6.0$, $-\text{NH}-\text{CO}$). $^{13}\text{C NMR}$ (150.9 MHz, DMSO): 28.33 ($(\text{CH}_3)_3$), 40.65 ($-\text{CH}_2-\text{N}$), 51.08 ($>\text{CH}-\text{N}$), 78.42 (O–C(CH_3) $_3$), 155.35 (N–CO–O), 171.18 (–COOH). IR (KBr) ν_{max} (cm^{-1}) 3348 m (NH); 1714 vs (C=O) acid; 1684 vs (C=O) carbamate; 1530 m (amide II); 1624 s (NH_2); 2978 m, 2932 m, 1366 m (CH_3). HRMS (ESI) calc for $\text{C}_{18}\text{H}_{15}\text{O}_4\text{N}_2$ [$\text{M} - \text{H}$] $^+$ 203.10373, found: 203.10366.

2-(5)-(tert-Butoxycarbonylamino)-3-azidopropionic Acid **32**

With the previously described azido transfer reaction employed for **11**, the reaction of **31** (12.7 g; 62.2 mmol), TEA (18.9 g; 186.6 mmol) and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (155 mg; 0.622 mmol) in a mixture of 100 ml of methanol and 50 ml of water with TfN_3 in dichloromethane gave **32**. TfN_3 was prepared by the reaction of NaN_3 (40.4 g; 622 mmol) and triflic anhydride (35.1 g; 124.4 mmol) in 100 ml of water and 100 ml of DCM. A bright yellow oil (9 g) was obtained after the isolation by flash chromatography on silica gel and was used as a crude product in the following step.

Copper(II) Complex of N^3 -tert-Butoxycarbonyl-L-ornithine **35**

Complex **35** was prepared according to the literature [58] by a method starting from L-ornithine. HCl **33** (8.4 g; 50 mmol), $\text{Cu}(\text{CH}_3\text{COO})_2 \cdot \text{H}_2\text{O}$ (5 g; 25 mmol), Boc_2O (12 g; 55 mmol) and 50 ml of 2 M NaOH. Yield 11.8 g (90%). Dark violet solid, >220 °C (decay). Because of the diamagnetism of copper, NMR spectra were not recorded. IR (KBr) ν_{max} (cm^{-1}) 1684 vs (C=O) carbamate; 1573 m, 1401 s (COO^-); 1620 vs (NH_2); 1522 s (amide II); 2979 m, 1392 s, 1367 (CH_3); 1174 vs ($\text{C}(\text{CH}_3)_3$). HRMS (ESI) calc for $\text{C}_{20}\text{H}_{39}\text{O}_8\text{N}_4\text{Cu}$ [$\text{M} + 1$] $^+$ 526.20584, found: 526.20605.

Copper(II) Complex of N^9 -tert-Butoxycarbonyl-L-lysine **36**

With the method employed for **35**, the complex of **36** was prepared from L-lysine-HCl **26** (9.13 g; 50 mmol), $\text{Cu}(\text{CH}_3\text{COO})_2 \cdot \text{H}_2\text{O}$ (5 g; 25 mmol), Boc_2O (12 g; 55 mmol) and 50 ml of 2 M NaOH. Yield 12.7 g (92%). Dark violet solid, >200 °C (decay). Because of the diamagnetism of copper, NMR spectra were not recorded. IR (KBr) ν_{max} (cm^{-1}) 1686 vs (C=O) carbamate; 1573 m, 1401 s (COO^-); 1623 vs (NH_2); 1522 m (amide II); 2979 m, 1392 s, 1367 (CH_3); 2934 m, 1456 m (CH_2); 1173 vs ($\text{C}(\text{CH}_3)_3$). HRMS (ESI) calc for $\text{C}_{22}\text{H}_{43}\text{O}_8\text{N}_4\text{Cu}$ [$\text{M} + 1$] $^+$ 554.23714, found: 554.23719.

5-(tert-Butoxycarbonylamino)-2-(5)-aminopentanoic Acid **37**

Protected acid **37** was prepared by the reaction of **35** (10.2 g; 19.4 mmol) and 8-quinolinol (7.3 g; 50.4 mmol), using the method described previously [56–59]. Yield 8.1 g (90%). White solid, m.p. 219–222 °C. $R_f = 0.63$ (isopropyl alcohol–concentrated aqueous ammonia–water 7 : 1 : 2). $[\alpha]_D^{20} = +16.7$ ($c = 0.222$; glacial acetic acid). $^1\text{H NMR}$ (500 MHz, DMSO): 1.37 (9H, s, $(\text{CH}_3)_3$), 1.42 (2H, m, $-\text{CH}_2-$), 1.51 and 1.67 (2H, 2 \times m, $-\text{CH}_2-$), 2.88 (2H, m, $-\text{CH}_2-\text{N}$), 3.07 (1H, br t, $J = 6.0$, $>\text{CH}-\text{N}$), 6.86 (1H, br t, $J = 5.4$, $-\text{NH}-\text{CO}$). $^{13}\text{C NMR}$ (125.7 MHz, DMSO): 26.14 ($-\text{CH}_2-$), 28.48 ($(\text{CH}_3)_3$), 28.80 ($-\text{CH}_2-$), 54.26 ($>\text{CH}-\text{N}$), 77.52 (O–C(CH_3) $_3$), 155.75 (N–CO–O), 169.75 (–COOH). IR (KBr) ν_{max} (cm^{-1}) 3362 s (NH); 2978 vs, 1393 s, 1366 s (CH_3); 1174 vs ($\text{C}(\text{CH}_3)_3$); 1688 vs (C=O) carbamate; 1527 s (amide II); ~1588 vs, 1407 vs (COO^-); 1588 s (NH_2). HRMS (ESI) calc for $\text{C}_{10}\text{H}_{21}\text{O}_4\text{N}_2$ [$\text{M} + 1$] $^+$ 233.14958, found: 233.14959.

6-(tert-Butoxycarbonylamino)-2-(5)-aminohexanoic Acid **38**

With the use of the same method as for **37**, protected acid **38** was prepared by the reaction of **36** (12.3 g; 22 mmol) and 8-quinolinol (8.3 g; 57.2 mmol). Yield 9.9 g (88%). White solid, m.p. 228–230 °C. $R_f = 0.64$ (isopropyl alcohol–concentrated aqueous ammonia–water 7 : 1 : 2). $[\alpha]_D^{20} = +4.8$ ($c = 0.270$; glacial acetic acid). $^1\text{H NMR}$ (600 MHz, DMSO): 1.28 (2H, m, $-\text{CH}_2-$), 1.34 (2H, m, $-\text{CH}_2-$), 1.37 (9H, s, $(\text{CH}_3)_3$), 1.53 and 1.68 (2H, 2 \times m, $-\text{CH}_2-$), 2.88 (2H, m, $-\text{CH}_2-\text{N}$), 3.06 (1H, dd, $J = 7.3$ and 5.0, $>\text{CH}-\text{N}$), 6.74 (1H, br t, $J = 5.5$, $-\text{NH}-\text{CO}$). $^{13}\text{C NMR}$ (150.9 MHz, DMSO): 22.72 ($-\text{CH}_2-$), 28.45 ($(\text{CH}_3)_3$), 29.41 ($-\text{CH}_2-$), 31.04 ($-\text{CH}_2-$), 39.95 ($-\text{CH}_2-\text{N}$), 54.34 ($>\text{CH}-\text{N}$), 77.51 (O–C(CH_3) $_3$), 155.70 (N–CO–O), 169.78 (–COOH). IR (KBr) ν_{max} (cm^{-1}) 3380 s (NH); 2978 vs, 1398 s, 1366 s (CH_3); 1178 vs ($\text{C}(\text{CH}_3)_3$); 1689 vs (C=O) carbamate; 1520 s (amide II); 1624 vs (NH_2); 1585 vs, 1407 vs (COO^-); 1588 s (NH_2). HRMS (ESI) calc for $\text{C}_{11}\text{H}_{22}\text{O}_4\text{N}_2\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 269.14718, found: 269.14719.

5-(tert-Butoxycarbonylamino)-2-(5)-(9)-fluorenylmethoxycarbonylamino)pentanoic Acid **39**

Compound **37** (5.9 g; 25.4 mmol) was placed in a 1 l round-bottom flask, equipped with a magnetic spin bar and dissolved in a solution of NaHCO_3 (4.3 g; 50.8 mmol) in 100 ml of water. The flask was immersed in an ice cooling bath and Fmoc-OSu (8.6 g; 25.4 mmol) in 100 ml of dioxane was added dropwise under vigorous stirring during 30 min. When the addition of Fmoc-OSu was complete, the reaction mixture was allowed to react for 1 h at 0 °C and then overnight at RT. Thereafter, 200 ml of water was added and followed by the dropwise addition of concentrated citric acid until pH ~ 2–3 was reached. The reaction mixture was extracted four times with 150 ml of ethyl acetate. The combined organic layers were washed twice with 150 ml of brine and twice with 150 ml of water and dried over Na_2SO_4 . The filtrate was evaporated and the resulting brown oil was subjected to flash chromatography on silica gel, using a linear gradient ethyl acetate in ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5. The yellow oil was triturated in a mixture of ethyl acetate–petroleum ether at -20 °C to afford the pure product. Yield 10.7 g (93%). White solid, m.p. 88–90 °C. $R_f = 0.69$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = -1.6$ ($c = 0.254$; DMF). $^1\text{H NMR}$ (500 MHz, DMSO): 1.37 (9H, s, $(\text{CH}_3)_3$), 1.43 (2H, m, $-\text{CH}_2-$), 1.56 and 1.71 (2H, 2 \times m, $-\text{CH}_2-$), 2.91 (2H, m, $-\text{CH}_2-\text{N}$), 3.90 (1H, ddd, $J = 9.1$, 8.2 and 4.8, $>\text{CH}-\text{N}$), 4.22 (1H, dd, $J = 7.1$ and 6.8, $>\text{CH}-$), 4.26 (1H, dd, $J = 10.0$ and 6.3, CO–O– $\text{CHaHb}-$), 4.28 (1H, dd, $J = 10.0$ and 7.1, CO–O– $\text{CHaHb}-$),

6.80 (1H, t, $J = 5.6$, -NH-CO-O), 7.52 (1H, br d, $J = 8.2$, -NH-CO-O), 7.33 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.72 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 12.70 (1H, br s, COOH). ^{13}C NMR (125.7 MHz, DMSO): 26.46 (-CH₂-), 28.47 ((CH₃)₃), 28.62 (-CH₂-), 40.00 (-CH₂-N), 54.13 (>CH-N), 65.78 (-CH₂-O-CO), 77.58 (O-C(CH₃)₃), 120.30(2), 125.50, 125.51, 127.28(2) and 127.83(2) (8x Ar=CH-), 140.91, 140.92, 144.02 and 144.08 (4x Ar>C=), 155.79 (N-CO-O), 156.24 (N-CO-O), 174.30 (-COOH). IR (KBr) ν_{max} (cm⁻¹) 3348 s (NH); 2976 vs, 1393 s, 1366 s (CH₃); 1169 vs (C(CH₃)₃); 1716 vs (C=O) acid; 1698 vs (C=O) carbamates; 1528 s (amide II); 3066 w, 3041 w, 1451 m, 1105 m, 759 m, 740 m (ring). HRMS (ESI) calc for C₂₅H₃₀O₆N₂Na [M + Na]⁺ 477.19961, found: 477.19964.

6-(tert-Butoxycarbonylamino)-2-(S)-(9-fluorenylmethyloxycarbonylamino)hexanoic Acid 40

Acid **40** was prepared by the reaction of **38** (9.3 g; 35.2 mmol), NaHCO₃ (5.9 g; 70.4 mmol) and Fmoc-Osu (11.1 g; 35.2 mmol), using the protocol described for **39**. Yield 15.6 g (95%). White solid, m.p. 125–127 °C. $R_f = 0.75$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = -8.5$ ($c = 0.272$; DMF). ^1H NMR (500 MHz, DMSO): 1.30 (2H, m, -CH₂-), 1.36 (9H, s, (CH₃)₃), 1.36 (2H, m, -CH₂-), 1.59 and 1.69 (2H, 2x m, -CH₂-), 2.90 (2H, m, -CH₂-N), 3.89 (1H, ddd, $J = 9.2$, 8.2 and 4.5, >CH-N), 4.22 (1H, m, >CH-), 4.27 (2H, m, CO-O-CH₂-), 6.79 (1H, br t, $J = 5.6$, -NH-CO-O), 7.54 (1H, br d, $J = 8.2$, -NH-CO-O), 7.33 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.73 (2H, m, Ar-H), 7.89 (2H, m, Ar-H). ^{13}C NMR (125.7 MHz, DMSO): 23.17 (-CH₂-), 28.49 ((CH₃)₃), 29.34 (-CH₂-), 30.83 (-CH₂-), 39.82 (-CH₂-N), 46.88 (>CH-), 54.21 (>CH-N), 65.78 (-CH₂-O-CO), 77.56 (O-C(CH₃)₃), 120.33, 120.34, 125.50, 125.52, 127.29(2) and 127.85(2) (8x Ar=CH-), 140.93, 140.94, 144.03 and 144.09 (4x Ar>C=), 155.80 (N-CO-O), 156.32 (N-CO-O), 174.40 (-COOH). IR (KBr) ν_{max} (cm⁻¹) 3391 s, 3369 s (NH); 2978 vs, 1393 s, 1367 s (CH₃); 2936 m, 1478 m (CH₂); 1173 vs (C(CH₃)₃); 1711 vs (C=O) acid; 1693 vs + br (C=O) carbamates; 1525 s (amide II); 3067 w, 3041 w, 1451 m, 1105 m, 760 m, 740 m (ring). HRMS (ESI) calc for C₂₆H₃₂O₆N₂Na [M + Na]⁺ 491.21526, found: 491.21527.

5-Azido-2-(S)-(9-fluorenylmethyloxycarbonylamino)pentanoic Acid 41

Compound **39** (10.4 g; 22.9 mmol) was treated with 20 ml of DCM, 20 ml of TFA and 2 ml of water. After 2 h stirring, volatile materials were evaporated to give a yellow oil, which was then suspended with NaHCO₃ (19.2 g; 229 mmol) and CuSO₄·5H₂O (57 mg; 0.229 mmol) in 100 ml of water and 150 ml of methanol. TfN₃, prepared by the reaction of NaN₃ (14.9 g; 229 mmol) and triflic anhydride (12.9 g; 45.8 mmol), was added dropwise to the slurry in 100 ml of DCM, followed by the work-up, which was the same as in the case of the diazotransfer reaction leading to the product **14**. Yield 8 g (92%). White solid, 127–128 °C. $R_f = 0.71$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = -6.7$ ($c = 0.254$; DMF). ^1H NMR (500 MHz, DMSO): 1.59 (2H, m, -CH₂-), 1.66 and 1.78 (2H, 2x m, -CH₂-), 3.34 (2H, m, -CH₂-N₃), 3.98 (1H, ddd, $J = 9.2$, 8.2 and 5.0, >CH-N), 4.23 (1H, br t, $J = 7.0$, >CH-), 4.30 (2H, m, CO-O-CH₂-), 7.33 (2H, m, Ar-H), 7.42 (2H, m, Ar-H), 7.70 (1H, br d, $J = 8.2$, -NH-CO-O), 7.73 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 12.66 (1H, br s, COOH). ^{13}C NMR (125.7 MHz, DMSO): 25.31 (-CH₂-), 28.21 (-CH₂-), 46.89 (>CH-), 50.45 (-CH₂-N₃), 53.57 (>CH-N), 65.81 (-CH₂-O-CO), 120.33, 120.35, 125.48, 125.50, 127.28(2) and 127.86(2) (8x Ar=CH-), 140.94, 140.96, 144.00 and 144.07 (4x Ar>C=), 156.38 (N-CO-O), 173.88 (-COOH). IR (KBr) ν_{max}

(cm⁻¹) 3391 s, 3339 s (NH); 2874 m (CH₂); 2096 vs (N₃); 1711 vs, (C=O) acid; 1729 vs, 1687 vs (C=O) carbamate; 1534 s (amide II); 3043 m, 1451 s, 1103 m, 1032 m, 759 s, 737 s, 622 s (ring). HRMS (ESI) calc for C₂₀H₂₀O₄N₄Na [M + Na]⁺ 403.13768, found: 403.13774.

6-Azido-2-(S)-(9-fluorenylmethyloxycarbonylamino)hexanoic Acid 42

With the procedure described for **41**, acid **42** was prepared starting from **40** (14.9 g; 31.8 mmol), NaHCO₃ (26.7 g; 318 mmol), CuSO₄·5H₂O (79 mg; 0.32 mmol), NaN₃ (20.7 g; 318 mmol) and triflic anhydride (17.9 g; 63.6 mmol). Yield 11.1 g (89%). White solid, 73–75 °C. $R_f = 0.74$ (ethyl acetate–acetone–methanol–water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = -15.7$ ($c = 0.261$; DMF). ^1H NMR (500 MHz, DMSO): 1.38 (2H, m, -CH₂-), 1.53 (2H, m, -CH₂-), 1.63 and 1.73 (2H, 2x m, -CH₂-), 3.32 (2H, t, $J = 6.8$, -CH₂-N₃), 3.94 (1H, ddd, $J = 9.5$, 8.0 and 4.6, >CH-N), 4.22 (1H, br t, $J = 7.0$, >CH-), 4.28 (2H, m, CO-O-CH₂-), 7.33 (2H, m, Ar-H), 7.42 (2H, m, Ar-H), 7.67 (1H, br d, $J = 8.0$, -NH-CO-O), 7.73 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 12.62 (1H, br s, COOH). ^{13}C NMR (125.7 MHz, DMSO): 23.14 (-CH₂-), 28.06 (-CH₂-), 30.50 (-CH₂-), 46.87 (>CH-), 50.73 (-CH₂-N₃), 53.88 (>CH-N), 65.81 (-CH₂-O-CO), 120.34, 120.35, 125.49, 125.52, 127.28(2) and 127.87(2) (8x Ar=CH-), 140.94, 140.95, 144.02 and 144.07 (4x Ar>C=), 156.40 (N-CO-O), 174.11 (-COOH). IR (KBr) ν_{max} (cm⁻¹) 3378 m (NH); 2937 s, 2886 m, 1477 m (CH₂); 2097 vs (N₃); 1717 vs, (C=O) acid; 1746 vs, 1698 vs (C=O) carbamate; 1525 s (amide II); 3065 m, 3043 m, 1451 s, 1102 m, 758 s, 739 s (ring). HRMS (ESI) calc for C₂₁H₂₂O₄N₄Na [M + Na]⁺ 417.15333, found: 417.15346.

General Protocol for the Manual Synthesis of Tripeptides 43–47

- Rink amide resin (400 μmol, loading 0.68 mmol/g) was placed in a 20 ml polypropylene syringe equipped with a polypropylene frit and swelled in 10 ml of DMF for 1 h.
- Fmoc group was cleaved by treatment with 20% piperidine/DMF (5 ml for 5 and 20 min), followed by five washings with 5 ml of DMF.
- Fmoc-Phe (619 mg; 1.6 mmol), HBTU (607 mg; 1.6 mmol) and DIPEA (418 μmol; 2.4 mmol) were added in 5 ml of DMF. The resin was stirred for 2 h and then washed five times with 5 ml of DMF. Step (iii) was repeated, and the resin was washed successively with DMF, MeOH, DCM and DMF (each solvent five times with 5 ml). Step (ii) was repeated. Step (iii) was repeated twice (Fmoc-Val was used in the case of tripeptide **43**; 543 mg; 1.6 mmol). Step (ii) was repeated.
- 14** (563 mg; 1.6 mmol) or **29** (586 mg; 1.6 mmol) or **41** (608 mg; 1.6 mmol) or **42** (630 mg; 1.6 mmol) with HBTU (607 mg; 1.6 mmol) and DIPEA (418 μmol; 2.4 mmol) in 5 ml of DMF was added, stirred for 2 h and then washed five times with 5 ml of DMF. Only in the case of the synthesis of **44** did couplings with **14** proceed for 5 and 18 h. Step (iv) was repeated and the resin successively washed with DMF, MeOH, DCM and DMF (each solvent five times with 5 ml). Step (ii) was repeated.
- Thereafter, 300 μl of Ac₂O and 300 μl of DIPEA, each in 1 ml of DMF, were added; the resin was stirred for 15 min and washed five times with 5 ml of DMF. Step (v) was repeated. Thereafter, the resin was transported to a small glass reactor equipped with a frit, rinsed with 50 ml of dichloromethane and dried overnight under deep vacuum. The resin was cleaved for 60 min with 5 ml of a cocktail of TFA/water/triisopropylsilane

(CAS 6485-79-6) (95/2.5/2.5). The cleavage step was repeated under the same conditions, the resin was washed with 10 ml of glacial acetic acid and all cleavage solutions and acetic acid were combined and evaporated under reduced pressure. The brown residue was then sonicated for 10 min with 10 ml diethyl ether in an ice cooling bath. The slurry was centrifuged for 10 min at 10 000 g, diethyl ether was decanted and crude amorphous white tripeptides were dried *in vacuo*. The purity of the prepared tripeptides was checked by RP-HPLC; analytical samples were isolated using the following gradient: $t = 0$ min (20% B), $t = 30$ min (100% B).

Analytical data for compounds **43–47** are provided in the supporting information.

CH₃COCO-Phe-Phe-CONH₂ **48**

Yield 49 mg (32%). Lyophilisate. [α]_D²⁰ = -14.3 ($c = 0.119$; DMSO). ¹H NMR (500 MHz, DMSO): 2.25 (3H, s, CH₃-CO), 2.81 (1H, dd, $J = 13.8$ and 9.1, -CHaHb-), 3.02 (1H, dd, $J = 13.8$ and 5.0, -CHaHb-), 2.89 (1H, dd, $J = 13.8$ and 9.3, -CHaHb-), 2.96 (1H, dd, $J = 13.8$ and 4.7, -CHaHb-), 4.44 (1H, ddd, $J = 9.3$, 8.7 and 4.7, >CH-N), 4.45 (1H, ddd, $J = 9.1$, 8.2 and 5.0, >CH-N), 7.12 and 7.42 (2H, 2× br s, CONH₂), 7.12–7.26 (10H, m, 2× C₆H₅), 8.21 (1H, d, $J = 8.2$, -NH-CO), 8.37 (1H, d, $J = 8.7$, -NH-CO). ¹³C NMR (125.7 MHz, DMSO): 24.98 (CH₃-CO), 37.23 (-CH₂-), 37.92 (-CH₂-), 54.02 (>CH-N), 54.36 (>CH-N), 126.52, 126.57, 128.28(4), 129.40(2) and 129.46(2) (10× Ar=CH-), 137.62 and 137.98 (2× Ar>C=), 160.61 (-CO-CO-N), 169.99 (N-CO-), 172.86 (-CONH₂), 196.73 (-CO-CO-N). IR (KBr) ν_{\max} (cm⁻¹) 3408 vs, 3320 (NH); 1725 m (C=O) ketone, 1665 s, 1640 vs (C=O) amides; 1524 m (amide II); 3086 w, 3065 w, 3029 w, 1498 w, 1455 w, 750 w, 703 w (ring); 1358 m (CH₃). HRMS (ESI) calc for C₂₁H₂₃O₄N₃Na [M + Na]⁺ 404.15808 found: 404.15815.

Acknowledgments

This work was supported by the Grant Agency of the Czech Republic (grant 14-17305S to J. J.), by the Czech Academy of Sciences (Research Project RVO:6138963, supporting the Institute of Organic Chemistry and Biochemistry) and partially by the Medical Research Council (grant MR/K000179/1).

References

- Tomoe CW, Christensen C, Meldal M. Peptidotriazoles on solid phase: 1,2,3-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* 2002; **67**: 3057–3064.
- Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chem.-Int. Edit.* 2002; **41**: 2596–2599.
- Himo F, Lovell T, Hilgraf R, Rostovtsev VV, Noodleman L, Sharpless KB, Fokin VV. Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. *J. Am. Chem. Soc.* 2005; **127**: 210–216.
- Bock VD, Hiemstra H, van Maarseveen JH. Cu-I-catalyzed alkyne-azide 'click' cycloadditions from a mechanistic and synthetic perspective. *Eur. J. Org. Chem.* 2006; 51–68.
- Ingale S, Dawson PE. On resin side-chain cyclization of complex peptides using CuAAC. *Org. Lett.* 2011; **13**: 2822–2825.
- Jagasia R, Holub JM, Bollinger M, Kirshenbaum K, Finn MG. Peptide cyclization and cyclodimerization by Cu-I-mediated azide-alkyne cycloaddition. *J. Org. Chem.* 2009; **74**: 2964–2974.
- Pedersen DS, Abell A. 1,2,3-Triazoles in peptidomimetic chemistry. *Eur. J. Org. Chem.* 2011; 2399–2411.
- Angell YL, Burgess K. Peptidomimetics via copper-catalyzed azide-alkyne cycloadditions. *Chem. Soc. Rev.* 2007; **36**: 1674–1689.
- Park JH, Waters ML. Positional effects of click cyclization on beta-hairpin structure, stability, and function. *Org. Biomol. Chem.* 2013; **11**: 69–77.
- Kawamoto SA, Coleska A, Ran X, Yi H, Yang CY, Wang SM. Design of triazole-stapled BCL9 alpha-helical peptides to target the beta-catenin/B-cell CLL/lymphoma 9 (BCL9) protein-protein interaction. *J. Med. Chem.* 2012; **55**: 1137–1146.
- Scrima M, Le Chevalier-Isaad A, Rovero P, Papini AM, Chorev M, D'Ursi AM. Cu-I-catalyzed azide-alkyne intramolecular *i*-to-*i* + 4) side-chain-to-side-chain cyclization promotes the formation of helix-like secondary structures. *Eur. J. Org. Chem.* 2010; 446–457.
- Holland-Nell K, Meldal M. Maintaining biological activity by using triazoles as disulfide bond mimetics. *Angew. Chem.-Int. Edit.* 2011; **50**: 5204–5206.
- Meldal M, Tornøe CW. Cu-catalyzed azide-alkyne cycloaddition. *Chem. Rev.* 2008; **108**: 2952–3015.
- Holub JM, Kirshenbaum K. Tricks with clicks: modification of peptidomimetic oligomers via copper-catalyzed azide-alkyne [3 + 2] cycloaddition. *Chem. Soc. Rev.* 2010; **39**: 1325–1337.
- Castro V, Rodriguez H, Abericio F. CuAAC: an efficient click chemistry reaction on solid phase. *ACS Comb. Sci.* 2016; **18**: 1–14.
- Nilsson BL, Kiessling LL, Raines RT. Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.* 2000; **2**: 1939–1941.
- Vikova J, Collinsova M, Kletvikova E, Budesinsky M, Kaplan V, Zakova L, Veverka V, Hexnerova R, Avino RJ, Strakova J, Selicharova I, Vanek V, Wright DW, Watson CJ, Turkenburg JP, Brzozowski AM, Jiracek J. Rational steering of insulin binding specificity by intra-chain chemical crosslinking. *Sci. Rep.* 2016; **6**: 12.
- Wei L, Lubell WD. Scope and limitations in the use of N-(Ph)serine-derived cyclic sulfamidates for amino acid synthesis. *Can. J. Chem.* 2001; **79**: 94–104.
- Sai Sudhir V, Phani Kumar NY, Nasir Baig RB, Chandrasekaran S. Facile entry into triazole fused heterocycles via sulfamidate derived azido-alkynes. *J. Org. Chem.* 2009; **74**: 7588–7591.
- Sudhir VS, Kumar NYP, Chandrasekaran S. Click chemistry inspired synthesis of ferrocene amino acids and other derivatives. *Tetrahedron* 2010; **66**: 1327–1334.
- Arnold LD, May RG, Vederas JC. Synthesis of optically pure alpha-amino acids via salts of alpha-amino-beta-propiolactone. *J. Am. Chem. Soc.* 1988; **110**: 2237–2241.
- Fujii M, Hidaka J. Nucleic acid analog peptide containing beta-aminoalanine modified with nucleobases. *Nucleos. Nucleot.* 1999; **18**: 1421–1422.
- Sun DQ, Jones V, Carson EI, Lee REB, Scherman MS, McNeil MR, Lee RE. Solid-phase synthesis and biological evaluation of a uridyl branched peptide urea library. *Bioorg. Med. Chem. Lett.* 2007; **17**: 6899–6904.
- Fujii M, Yoshida K, Hidaka J, Ohtsu T. Nucleic acid analog peptide (NAAP). 2. Syntheses and properties of novel DNA analog peptides containing nucleobase linked beta-aminoalanine. *Bioorg. Med. Chem. Lett.* 1997; **7**: 637–640.
- Fujii M, Yoshida K, Hidaka J, Ohtsu T. Hybridization properties of nucleic acid analogs containing beta-aminoalanine modified with nucleobases. *Chem. Commun.* 1998; 717–718.
- Rosenberg SH, Spina KP, Woods KW, Polakowski J, Martin DL, Yao ZL, Stein HH, Cohen J, Barlow JL, Egan DA, Tricarico KA, Baker WR, Kleinert HD. Studies directed toward the design of orally active renin inhibitors. 1. Some factors influencing the absorption of small peptides. *J. Med. Chem.* 1993; **36**: 449–459.
- Kogan TP, Rawson TE. The synthesis of chiral 3-oxo-6-(phenylmethoxy)-carbonyl-2-piperazineacetic acid-esters designed for the presentation of an aspartic-acid side-chain – a subsequent novel Friedel crafts reaction. *Tetrahedron Lett.* 1992; **33**: 7089–7092.
- Pickersgill IF, Rapoport H. Preparation of functionalized, conformationally constrained DTPA analogues from L- or D-serine and trans-4-hydroxy-L-proline. Hydroxymethyl substituents on the central acetic acid and on the backbone. *J. Org. Chem.* 2000; **65**: 4048–4057.
- Gajewski M, Seaver B, Esslinger CS. Design, synthesis, and biological activity of novel triazole amino acids used to probe binding interactions between ligand and neutral amino acid transport protein SN1. *Bioorg. Med. Chem. Lett.* 2007; **17**: 4163–4166.
- Colombo R, Mingozzi M, Belvisi L, Arosio D, Piarulli U, Carenini N, Perego P, Zaffaroni N, De Cesare M, Castiglioni V, Scanziani E, Gennari C. Synthesis and biological evaluation (*in vitro* and *in vivo*) of cyclic arginine-glycine-aspartate (RGD) peptidomimetic-paclitaxel

- conjugates targeting integrin alpha(v)beta(3). *J. Med. Chem.* 2012; **55**: 10460–10474.
- 31 Otsuka M, Kittaka A, Iimori T, Yamashita H, Kobayashi S, Ohno M. Synthetic studies on an antitumor antibiotic, bleomycin. 12. Preparation of an L-2,3-diaminopropionic acid-synthetic intermediate. *Chem. Pharm. Bull.* 1985; **33**: 509–514.
- 32 Stanley NJ, Pedersen DS, Nielsen B, Kvist T, Mathiesen JM, Brauner-Osborne H, Taylor DK, Abell AD. 1,2,3-Triazolyl amino acids as AMPA receptor ligands. *Bioorg. Med. Chem. Lett.* 2010; **20**: 7512–7515.
- 33 Boger DL, Honda T, Menezes RF, Colletti SL, Dang Q, Yang WJ. Total syntheses of (+)-P-3A, epi-(–)-P-3A, and (–)-desacetamido P-3A. *J. Am. Chem. Soc.* 1994; **116**: 82–92.
- 34 Mukai S, Flematt GR, Byrne LT, Besant PG, Attwood PV, Piggott MJ. Stable triazolylphosphonate analogues of phosphohistidine. *Amino Acids* 2012; **43**: 857–874.
- 35 Panda G, Rao NV. A short synthetic approach to chiral serine azido derivatives. *Synlett* 2004; 714–716. DOI: 10.1055/s-2004-817770.
- 36 Jeong JM, Shetty D, Lee DS, Chung JK, Lee MC, Jeong J, Dineswi S, Lee D, US2012029177-A1, 2011.
- 37 Friscount F, Fahrni CJ, Boons GJ. A fluorogenic probe for the catalyst-free detection of azide-tagged molecules. *J. Am. Chem. Soc.* 2012; **134**: 18809–18815.
- 38 Zhong M, Hanan EJ, Shen W, Bui M, Arkin MR, Barr KJ, Evanchik MJ, Hoch U, Hyde J, Martell JR, Oslob JD, Paulvannan K, Prabhu S, Silverman JA, Wright J, Yu CH, Zhu JA, Flanagan WM. Structure–activity relationship (SAR) of the alpha-amino acid residue of potent tetrahydroisoquinoline (THIQ)-derived LFA-1/ICAM-1 antagonists. *Bioorg. Med. Chem. Lett.* 2011; **21**: 307–310.
- 39 Zou Y, Fahmi NE, Vialas C, Miller GM, Hecht SM. Total synthesis of deamido bleomycin A(2), the major catabolite of the antitumor agent bleomycin. *J. Am. Chem. Soc.* 2002; **124**: 9476–9488.
- 40 Zhang XJ, Krishnamurthy R. Mapping the landscape of potentially primordial informational oligomers: oligo-dipeptides tagged with orotic acid derivatives as recognition elements. *Angew. Chem.-Int. Edit.* 2009; **48**: 8124–8128.
- 41 Lau YH, Spring DR. Efficient synthesis of Fmoc-protected azido amino acids. *Synlett* 2011; 1917–1919. DOI: 10.1055/s-0030-1260950.
- 42 Akaji K, Aimoto S. Synthesis of MEN11420, a glycosylated bicyclic peptide, by intramolecular double cyclization using a chloroimidazolium coupling reagent. *Tetrahedron* 2001; **57**: 1749–1755.
- 43 Hirschmann R, Yao WQ, Arison B, Maechler L, Rosegay A, Sprengeler PA, Smith AB. Synthesis of the first tricyclic homodetic peptide. Use of coordinated orthogonal deprotection to achieve directed ring closure. *Tetrahedron* 1998; **54**: 7179–7202.
- 44 Barghash RF, Massi A, Dondoni A. Synthesis of thiourea-tethered C-glycosyl amino acids via isothiocyanate-amine coupling. *Org. Biomol. Chem.* 2009; **7**: 3319–3330.
- 45 Zhang LH, Kauffman GS, Pesti JA, Yin JG. Rearrangement of N-alpha-protected L-asparagines with iodobenzene diacetate. A practical route to beta-amino-L-alanine derivatives. *J. Org. Chem.* 1997; **62**: 6918–6920.
- 46 Millward SW, Henning RK, Kwong GA, Pitram S, Agnew HD, Deyle KM, Nag A, Hein J, Lee SS, Lim J, Pfeilsticker JA, Sharpless KB, Heath JR. Iterative *in situ* click chemistry assembles a branched capture agent and allosteric inhibitor for Akt1. *J. Am. Chem. Soc.* 2011; **133**: 18280–18288.
- 47 Ghosh PS, Hamilton AD. Noncovalent template-assisted mimicry of multiloop protein surfaces: assembling discontinuous and functional domains. *J. Am. Chem. Soc.* 2012; **134**: 13208–13211.
- 48 Pehere AD, Abell AD. New beta-strand templates constrained by Huisgen cycloaddition. *Org. Lett.* 2012; **14**: 1330–1333.
- 49 Oh KI, Lee JH, Joo C, Han H, Cho M. Beta-azidoalanine as an IR probe: application to amyloid A beta(16–22) aggregation. *J. Phys. Chem. B* 2008; **112**: 10352–10357.
- 50 Miller N, Williams GM, Brimble MA. Synthesis of fish antifreeze neoglycopeptides using microwave-assisted 'click chemistry'. *Org. Lett.* 2009; **11**: 2409–2412.
- 51 Roice M, Johannsen I, Meldal M. High capacity poly(ethylene glycol) based amino polymers for peptide and organic synthesis. *QSAR Comb. Sci.* 2004; **23**: 662–673.
- 52 Thuriereau C, Janiak P, Krantzig S, Guyard C, Pillon A, Kucharczyk N, Vilaine JP, Fauchere JL. A new somatostatin analog with optimized ring size inhibits neointima formation induced by balloon injury in rats without altering growth-hormone release. *Eur. J. Med. Chem.* 1995; **30**: 115–122.
- 53 Sakura N, Itoh T, Uchida Y, Ohki K, Okimura K, Chiba K, Sato Y, Sawanishi H. The contribution of the N-terminal structure of polymyxin B peptides to antimicrobial and lipopolysaccharide binding activity. *B. Chem. Soc. Jpn.* 2004; **77**: 1915–1924.
- 54 Papeo G, Giordano P, Brasca MG, Buzzo F, Caronni D, Ciprandi F, Mongelli N, Veronesi M, Vulpetti A, Dalvit C. Polyfluorinated amino acids for sensitive F-19 NMR-based screening and kinetic measurements. *J. Am. Chem. Soc.* 2007; **129**: 5665–5672.
- 55 Roth S, Thomas NR. A concise route to L-azidoamino acids: L-azidoalanine, L-azidohomoalanine and L-azidonorvaline. *Synlett* 2010; 607–609. DOI: 10.1038/nprot.2010.164.
- 56 Masiukiewicz E, Wiejak S, Rzeszotarska B. Scalable syntheses of N-alpha-benzyloxycarbonyl-L-ornithine and of N-alpha-(9-fluorenylmethoxy) carbonyl-L-ornithine. *Org. Proc. Int.* 2002; **34**: 531–537.
- 57 Wiejak S, Masiukiewicz E, Rzeszotarska B. A large scale synthesis of mono- and di-urethane derivatives of lysine. *Chem. Pharm. Bull.* 1999; **47**: 1489–1490.
- 58 Wiejak S, Masiukiewicz E, Rzeszotarska B. Improved scalable syntheses of mono- and bis-urethane derivatives of ornithine. *Chem. Pharm. Bull.* 2001; **49**: 1189–1191.
- 59 Bayryamov SG, Vassilev NG, Petkov DD. The two pathways for effective orthogonal protection of L-ornithine, for amino acylation of 5'-O-pivaloyl nucleosides, describe the general and important role for the successful imitation, during the synthesis of the model substrates for the ribosomal mimic reaction. *Protein Pept. Lett.* 2010; **17**: 889–898.
- 60 Strazzolini P, Scuccato M, Giumanini AG. Deprotection of t-butyl esters of amino acid derivatives by nitric acid in dichloromethane. *Tetrahedron* 2000; **56**: 3625–3633.
- 61 Chevallet P, Garrouste P, Malawska B, Martinez J. Facile synthesis of tert-butyl ester of N-protected amino-acids with tert-butyl bromide. *Tetrahedron Lett.* 1993; **34**: 7409–7412.
- 62 Fernandez S, Crocamo N, Inceri M, Giglio J, Scarone L, Rey A. Preparation and preliminary bioevaluation of a ^{99m}Tc(CO)₃-glucose derivative prepared by a click chemistry route. *J. Label. Compd. Radiopharm.* 2012; **55**: 274–280.
- 63 Nishiyama K, Karigomi H. Reaction of trimethylsilyl azide with organic halides. *Chem. Lett.* 1982; **11**: 1477–1478.
- 64 Photaki I. Transformation of serine to cysteine. β -Elimination reactions in serine derivatives. *J. Am. Chem. Soc.* 1963; **85**: 1123–1126.
- 65 Johansson H, Pedersen DS. Azide- and alkyne-derivatized alpha-amino acids. *Eur. J. Org. Chem.* 2012; 4267–4281.
- 66 Photaki I, Bardakos V. Transformation of L-serine peptides to L-cysteine peptides. *J. Am. Chem. Soc.* 1965; **87**: 3489–3492.
- 67 Han JR, Lian JT, Tian X, Zhou SW, Zhen XL, Liu SX. Total synthesis of micromide: a marine natural product. *Eur. J. Org. Chem.* 2014; **7232**–7238.
- 68 Inamoto A, Ogasawara K, Omata K, Kabuto K, Sasaki Y. Samarium(III)-propylene diamine tetraacetate complex: a water-soluble chiral shift reagent for use in high-field NMR. *Org. Lett.* 2000; **2**: 3543–3545.
- 69 Hrubal L, Budesinsky M, Picha J, Jiracek J, Vanek V. Simplified syntheses of the water-soluble chiral shift reagents Sm-(R)-pdta and Sm-(S)-pdta. *Tetrahedron Lett.* 2013; **54**: 6296–6297.
- 70 Fields GB, Noble RL. Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbonyl amino-acids. *Int. J. Pept. Prot. Res.* 1990; **35**: 161–214.
- 71 Angelini G, Speranza M. Gas-phase acid-induced nucleophilic displacement-reactions. 5. Quantitative-evaluation of neighboring-group participation in bifunctional compounds. *J. Am. Chem. Soc.* 1981; **103**: 3800–3806.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

Supporting Information

for

Optimized syntheses of Fmoc azido amino acids for the preparation of azidopeptides

Jan Pícha, Miloš Buděšínský, Kateřina Macháčková, Michaela Collinsová and Jiří Jiráček*

*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, v.v.i.,
Flemingovo nám. 2, 166 10 Prague 6, Czech Republic
E-mail: jiracek@uochb.cas.cz*

Table of contents:

Experimental procedures and analytical data for compound 3	Page 2
Experimental procedures and analytical data for compound 4	Page 2
Experimental procedures and analytical data for compound 5	Page 3
Experimental procedures and analytical data for compound 20	Page 4
Experimental procedures and analytical data for compound 25	Page 4
Experimental procedures and analytical data for compound 26	Page 5
Experimental procedures and analytical data for compound 30	Page 6
Analytical data for compound 43	Page 6
Analytical data for compound 44	Page 7
Analytical data for compound 45	Page 7
Analytical data for compound 46	Page 8
Analytical data for compound 47	Page 8

2-(*S*)-(Benzyloxycarbonylamino)-3-hydroxypropanoic acid **3**

L-Serine **1** (10.5 g; 0.1 mol; $[\alpha]_D^{20} = +13$, $c = 5$, 5M HCl) was dissolved in 200 ml of 10 % aqueous solution of sodium carbonate in a 1 L round-bottom flask equipped with magnetic spin bar. The flask was immersed in an ice cooling bath, and benzyl chloroformate (1.1 eq.; 18.8 g; 0.11 mol) in 100 ml of dioxane was added dropwise during a period of 30 minutes under vigorous stirring. When the addition of ZCl was completed, the reaction mixture was allowed to react for 1 h at 0 °C and then overnight at room temperature. The reaction mixture was transferred to the separatory funnel, 500 ml of water was added and the mixture extracted by 200 ml of diethyl ether. The separated aqueous-dioxane layer was poured back into the reaction flask. During cooling, conc. HCl was added dropwise under vigorously stirring until pH \approx 1 was reached. A cloudy solution was extracted with 3 x 400 ml ethyl acetate; combined fractions were washed 2 x with 150 ml of brine and then dried with solid Na₂SO₄. The drying agent was filtered off and the filtrate evaporated under reduced pressure. Crystallization of the product from a mixture of ethyl acetate and petroleum ether afforded colorless crystals. Yield 18.3 g (77%), m.p. 114-116 °C. $R_f = 0.70$ (ethyl acetate-acetone-methanol-water 6 : 1 : 1 : 0.5). $[\alpha]_D^{20} = +6.6$ ($c = 1.089$; CH₃COOH). ¹H NMR (600 MHz, DMSO): 3.65 (1H, dd, $J = 11.2$ and 5.6, HO-CH_aH_b-), 3.67 (1H, dd, $J = 11.2$ and 4.5, O-CH_aH_b-), 4.06 (1H, ddd, $J = 8.2$, 5.6 and 4.5, >CH-N), 5.03 (1H, d, $J = 12.5$, CO-O-CH_aH_b-), 5.05 (1H, d, $J = 12.5$, CO-O-CH_aH_b-), 7.29 (1H, d, $J = 8.2$, -NH-CO), 7.31 – 7.37 (5H, m, C₆H₅); ¹³C NMR (150.9 MHz, DMSO): 56.84 (>CH-N); 61.52 (-CH₂-O), 65.65 (-CH₂-O-CO), 127.92(2), 128.00 and 128.54(2) (5x Ar=CH-), 137.17 (Ar >C=), 156.22 (N-CO-O), 172.32 (O-CO-). IR (KBr) ν_{\max} cm⁻¹ 3436 s (OH); 3336 s, 3319 s (NH); 1747 vs (C=O) acid; 1690 vs (C=O) carbamate; 1534 vs (amide II); 3061 m, 3029 m, 1083 m, 1060 s, 1029 s, 697 s (ring). HRMS (ESI) calc for C₁₁H₁₂O₅N [M+H]⁺ 238.07210, found: 238.07195.

tert-Butyl 2-(*S*)-(benzyloxycarbonylamino)-3-hydroxypropanoate **4**

The solution Z-L-Ser **3** (18 g; 75.3 mmol) in 400 ml of *N,N*-dimethylacetamide and 200 ml of *tert*-butyl bromide was mixed with finely ground anhydrous potassium carbonate (69.1 g; 0.5 mol) and TEBAAC (11.4 g; 0.05 mol). The reaction mixture was vigorously stirred and heated at 60 °C overnight. After cooling, the content of the flask was poured on to 500 g of crushed ice. The bright brown solution was extracted 4 x with 200 ml of ethyl acetate. Combined organic layers were washed 2 x with 200 ml of water and dried over sodium sulfate. The drying agent was filtered off and the filtrate was evaporated under reduced pressure to afford 22 g of crude

oily material, which was purified by flash chromatography on silica gel, using a linear gradient of ethyl acetate in toluene. The required product was a white solid and an analytical sample was prepared by crystallization from ethyl acetate and petroleum ether. Yield 14.7 g (66%), m.p. 91-92 °C. $R_f = 0.69$ (toluene-ethyl acetate 50 : 50). $[\alpha]_D^{20} = -14.6$ (c = 1.006; C₂H₅OH). ¹H NMR (600 MHz, DMSO): 1.39 (9H, s, (CH₃)₃), 3.62 (1H, dt, $J = 11.0, 6.6$ and 6.0 , HO-CH_aH_b-), 3.65 (1H, ddd, $J = 11.0, 6.6$ and 5.3 , O-CH_aH_b-), 4.00 (1H, ddd, $J = 6.6, 6.0$ and 4.0 , >CH-N), 4.85 (1H, dd, $J = 6.6$ and 5.3 , -OH), 5.04 (2H, s, CO-O-CH₂-), 7.34 (1H, d, $J = 4.4$, -NH-CO), 7.31 – 7.37 (5H, m, C₆H₅); ¹³C NMR (150.9 MHz, DMSO): 27.85 ((CH₃)₃), 57.49 (>CH-N), 61.56 (-CH₂-O), 65.63 (-CH₂-O-CO), 80.77 (O-C(CH₃)₃), 127.91(2), 127.99 and 128.51(2) (5x Ar =CH-), 137.17 (Ar >C=), 156.19 (N-CO-O), 169.99 (O-CO-). IR (KBr) ν_{\max} cm⁻¹ 3410 s (OH); 3266 s (NH); 1727 vs (C=O) ester; 1714 vs (C=O) carbamate; 1555 s (amide II); 1456 s, 1395 s, 1367 s (CH₃); 1240 vs, 1159 vs (C-O-C); 3064 m, 3037 m, 1029 s, 700 s (ring). HRMS (ESI) calc for C₁₅H₂₁O₅NNa [M+Na]⁺ 318.13119, found: 318.13101.

tert-Butyl 2-(*S*)-(9-fluorenylmethoxycarbonylamino)-3-hydroxypropanoate **5**

Z-L-Ser-*O**t*Bu **4** (15.2 g; 51.5 mmol) in a glass pressure bottle was dissolved in 300 ml of methanol and 500 mg of 10 % Pd/C was added. The mixture was vigorously stirred and allowed to react under the atmosphere of hydrogen (15 psi) at rt overnight. TLC analysis revealed (toluene-ethyl acetate 50:50) that the starting compound had completely disappeared. The catalyst was filtered off through celite and the celite was washed with 300 ml of methanol. The filtrate was evaporated *in vacuo* to give 8.1 g of yellow residue, which was immediately dissolved in 150 ml of saturated solution of NaHCO₃. The flask was placed in the ice bath and Fmoc-OSu (17 g; 50.3 mmol) in 150 ml dioxane was added dropwise under stirring. After the addition of the total amount of acylation agent, stirring continued for 1 h at 0 °C and then overnight at room temperature. Thereafter, 250 ml of water was added and the reaction mixture was transferred to the separatory funnel and extracted 4 x with 150 ml of ethyl acetate. Combined organic layers were washed consecutively, once with 150 ml water, twice with 150 ml of brine and dried over anhydrous Na₂SO₄. Evaporation of the filtrate under reduced pressure furnished a colorless oil, which solidified upon standing at 5 °C. An analytical sample was prepared by crystallization from a mixture of diethyl ether-petroleum ether. Colorless solid. Yield 14.6 g (74% over two steps). m.p. 124-125 °C. $R_f = 0.59$ (toluene-ethyl acetate 50 : 50). $[\alpha]_D^{20} = +7.3$ (c = 0.998; CHCl₃). ¹H NMR (600 MHz, DMSO): 1.39 (9H, s, (CH₃)₃), 4.01 (1H, ddd, $J = 8.1, 5.9$ and 4.5 , >CH-N), 4.23 (1H, dd, $J = 7.1$ and 6.8 , >CH-), 4.29 (1H, dd, $J = 10.5$

and 6.8, CO–O–CH_aH_b–), 4.33 (1H, dd, $J = 10.5$ and 7.1 , CO–O–CH_aH_b–), 4.65 (2H, m, –O–CH₂–), 4.88 (1H, t, $J = 5.9$, –OH), 7.43 (1H, d, $J = 8.1$, –NH–CO), 7.33 (2H, m, Ar-H), 7.42 (2H, m, Ar-H), 7.74 (2H, m, Ar-H), 7.89 (2H, m, Ar-H). ¹³C NMR (150.9 MHz, DMSO): 27.87 ((CH₃)₃), 46.82 (>CH–), 57.48 (>CH–N); 61.57 (–CH₂–O), 65.88 (–CH₂–O–CO), 80.77 (O–C(CH₃)₃), 120.29(2), 125.45, 125.42, 127.25(2) and 127.82(2) (8x Ar =CH–), 140.92(2) 143.99 and 144.02 (4x Ar >C=), 156.24 (N–CO–O), 170.03 (O–CO–). IR (KBr) ν_{\max} cm⁻¹ 3403 m (OH); 1737 vs (C=O) ester; 1682 vs C=O (carbamate); 1540 vs (amide II); 1261 s, 1156 s (C–O–C); 3068 w, 3041 w, 1452 m, 1033 m, 764 s, 741 s (ring); 2978 m, 1395 m, 1371 m (CH₃). HRMS (ESI) calc for C₂₂H₂₆O₅N [M+H]⁺ 384.18055, found: 384.18058.

L-Val-O t Bu.TsOH **20**

L-Valine **19** (5.9 g; 50 mmol) was suspended in 70 ml of *tert*-butyl acetate, and 6.5 ml of 70 % HClO₄ was slowly added under ice-cooling. The ice bath was removed and the reaction mixture was allowed to react at rt overnight. The reaction mixture was poured into 200 ml of water; the separated aqueous layer was basified with K₂CO₃ until pH ~ 10 was reached and then extracted 3 x with 100 ml of diethyl ether. The combined organic phases were dried over Na₂SO₄, filtered and carefully evaporated to afford crude L-Val-O t Bu (7.6 g; 43.8 mmol). The slightly yellow oil was dissolved in 20 ml of methanol and the solution of TsOH.H₂O (8.4 g; 43.8 mmol) in 20 ml methanol was added in one portion. The solvent was evaporated and the product was crystallized from a mixture of *i*-PrOH-diethyl ether-hexane. Yield 9.3 g (63%). White solid, m. p. 138-140°C. $[\alpha]_D^{20} = +11.5$ (c = 0.374 ; CH₃OH). ¹H NMR (600 MHz, DMSO): 0.95 (3H, d, $J = 7.0$, CH₃), 0.98 (3H, d, $J = 7.0$, CH₃), 1.46 (9H, s, (CH₃)₃), 2.12 (1H, m, >CH–), 2.29 (3H, br s, Ar–CH₃), 3.80 (1H, d, $J = 4.2$, >CH–N), 7.12 (2H, m, Ar-H), 7.48 (2H, m, Ar-H). ¹³C NMR (150.9 MHz, DMSO): 17.60 (CH₃), 18.39 (CH₃), 21.00 (Ar–CH₃), 27.80 ((CH₃)₃), 29.60 (>CH–), 57.75 (>CH–N), 83.25 (O–C(CH₃)₃), 125.70(2) and 128.30(2) (4x Ar =CH–), 137.92 and 145.77 (2x Ar >C=), 168.20 (–O–CO–). IR (KBr) ν_{\max} cm⁻¹ 2971 m, 2934 m, 1474 m, 1369 m (CH₃); 1742 vs (C=O); 1254 s (C–O); 1618 m, 1521 m (NH₃⁺); 3061 w, 3040 w, 1498 m, 1106 m, 1011 s, 849 m, 813 m, 676 s (ring); 1227 s, 1192 s, 1163 s, 1035 s (SO₃⁻). HRMS (ESI) calc for C₉H₂₀O₂N [M]⁺ 174.14886, found: 174.14871.

(9-Fluorenylmethyloxycarbonyl)asparagine **25**

L-Asparagine **23** (5 g; 37.8 mmol, $[\alpha]_D^{20} = +27^\circ - 31^\circ$, c = 13.2 in 1M HCl) was placed in a 1 L round-bottom flask, equipped with a magnetic spin bar, and dissolved in a solution of Na₂CO₃

(4 g; 37.8 mmol) in 50 ml of water. The flask was immersed in an ice cooling bath and Fmoc-OSu (12.8 g; 37.8 mol) in 70 ml of dioxane was added dropwise under vigorous stirring during a period of 30 min. When the addition of Fmoc-Osu was complete, the reaction mixture (a dense slurry) was allowed to react for 1 h at 0 °C and then overnight at room temperature. The reaction mixture was cooled again in an ice bath and concentrated HCl was added dropwise until pH ~ 0-1 was reached. The crystals were filtered off in a Büchner funnel and rinsed with 100 ml of chilled water. The filter cake was frozen and dried by lyophilization. The final purification was carried out by dissolving the solid in 100 ml of hot chloroform (60 °C). After cooling, the crystals were filtered off and washed with petroleum ether. Yield 11.6 g (87%). White solid, m.p. 181-183°C. $R_f = 0.76$ (ethyl acetate-acetone-methanol-water 4 : 1 : 1: 1). $[\alpha]_D^{20} = -11.1$ ($c = 0.350$; DMF). $^1\text{H NMR}$ (600 MHz, DMSO): 2.46 (1H, dd, $J = 15.5$ and 8.0 , $-\text{CHaHb}-\text{CO}$), 2.56 (1H, dd, $J = 15.5$ and 5.2 , $-\text{CHaHb}-\text{CO}$), 4.22 (1H, m, $>\text{CH}-$), 4.27 (2H, m, $-\text{CH}_2-\text{O}$), 6.92 and 7.35 (2H, 2x br s, $-\text{CO}-\text{NH}_2$), 7.53 (1H, d, $J = 8.4$, $-\text{NH}-\text{CO}$), 7.33 (2H, m, Ar-H), 7.42 (2H, m, Ar-H), 7.71 (2H, m, Ar-H), 7.89 (2H, m, Ar-H), 12.59 (1H, br s, COOH). $^{13}\text{C NMR}$ (150.9 MHz, DMSO): 36.91 ($-\text{CH}_2-$), 46.78 ($>\text{CH}-$), 50.74 ($>\text{CH}-\text{N}$), 65.87 ($-\text{CH}_2-\text{O}-\text{CO}$), 120.29(2), 125.43, 125.45, 127.28(2) and 127.82(2) (8x Ar $=\text{CH}-$), 140.88(2), 143.97 and 144.00 (4x Ar $>\text{C}=\text{C}$), 155.98 (N-CO-O), 171.36 (N-CO-), 173.36 (-COOH). IR (KBr) ν_{max} cm^{-1} 3405 s (NH); 3432 s, 3324 s (NH₂); 1735 vs (C=O) acid; 1697 vs (C=O) carbamate; 1657 vs (C=O) amide; 1541 s (amide II); 3065 m, 1478 m, 1211 s, 758 m, 739 s (ring). HRMS (ESI) calc for C₁₉H₁₈O₅N₂Na[M+Na]⁺ 377.11079, found: 377.11084.

(9-Fluorenylmethyloxycarbonyl)glutamine **26**

Compound **18** was prepared by the reaction of L-Gln **24** (5 g; 34.2 mmol, $[\alpha]_D^{20} = +33^\circ$, $c = 5\%$ in 5M HCl), and Fmoc-OSu (11.5 g; 34.2 mmol) by the method previously used for **25**. Yield 11.7 g (93%). $R_f = 0.76$ (ethyl acetate-acetone-methanol-water 4 : 1 : 1: 1). $[\alpha]_D^{20} = -13.6$ ($c = 0.286$; DMF). $^1\text{H NMR}$ (500 MHz, DMSO): 1.76 and 1.97 (2H, 2x m, $-\text{CH}_2-$), 2.16 (2H, t, $J = 7.7$, $-\text{CH}_2-\text{CO}$), 3.94 (1H, ddd, $J = 9.5$, 7.9 and 4.7 , $>\text{CH}-\text{N}$), 4.21 (1H, m, $>\text{CH}-$), 4.26 (2H, m, $-\text{CH}_2-\text{O}$), 6.79 and 7.33 (2H, 2x br s, $-\text{CO}-\text{NH}_2$), 7.68 (1H, d, $J = 7.9$, $-\text{NH}-\text{CO}$), 7.33 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.72 (2H, m, Ar-H), 7.88 (2H, m, Ar-H). $^{13}\text{C NMR}$ (125.7 MHz, DMSO): 26.82 ($-\text{CH}_2-$), 31.68 ($-\text{CH}_2-$), 46.91 ($>\text{CH}-$), 53.77 ($>\text{CH}-\text{N}$), 65.98 ($-\text{CH}_2-\text{O}-\text{CO}$), 120.42(2), 125.60(2), 127.42(2) and 127.98(2) (8x Ar $=\text{CH}-$), 141.00(2) and 144.09(2) (4x Ar $>\text{C}=\text{C}$), 156.48 (N-CO-O), 173.90 ($-\text{CO}-\text{NH}_2$), 174.07 (-COOH). IR (KBr)

ν_{\max} cm^{-1} 3429 m, 3335 w (NH); 1724 vs (C=O) acid; 1698 vs (C=O) carbamate; 1658 vs (C=O) amide; 1532 m (amide II); 3066 w, 3040 w, 3021 w, 1670 w, 1450 m, 1215 m, 751 m, 740 m (ring). HRMS (ESI) calc for $\text{C}_{20}\text{H}_{20}\text{O}_5\text{N}_2\text{Na}[\text{M}+\text{Na}]^+$ 391.12644, found: 377.12656.

(*tert*-Butoxycarbonyl)asparagine **30**

Boc₂O (25.7 g; 0.1176 mol) in 150 ml of dioxane was added dropwise to an ice-cooled solution of L-asparagine **23** (14.9 g; 0.112 mol) and Na₂CO₃ (11.9 g; 0.112 mol) in 150 ml of water. The reaction mixture was stirred for 1 h at 0 °C and then at rt overnight. The volume of the reaction mixture was reduced to half; the slurry was cooled in an ice bath and pH was adjusted by the addition of an aqueous solution of conc. KHSO₄ to 2-3. The crystals were filtered off in a Büchner funnel and rinsed with 100 ml of chilled water. The filter cake was lyophilized and the pure product was gained by crystallization from methanol-ethyl acetate-petroleum ether mixture.

Intermediate **30** (18.9 g; 81.4 mmol) was suspended in a mixture of 90 ml of acetonitrile, 90 ml of ethyl acetate and 45 ml of water. Thereafter, (diacetoxyiodo)benzene (31.4 g; 97.7 mmol) was added in 5 portions within 15 minutes. Ten minutes after the addition of the total amount of (diacetoxyiodo) benzene, the slurry turned clear, followed by rapid precipitation of the crude product. The filter cake was washed out with 200 ml of chilled ethyl acetate and no additional purification was needed. Yield 12.7 g (77%). White solid, m.p. 209-211 °C. $R_f = 0.40$ (ethyl acetate-acetone-methanol-water 4 : 1 : 1 : 1). $[\alpha]_D^{20} = -5.3$ (c = 0.318 ; acetic acid). ¹H NMR (600 MHz, DMSO): 1.39 (9H, s, (CH₃)₃), 2.77 (1H, dd, $J = 11.9$ and 9.1 , -CH_aH_b-N), 3.01 (1H, dd, $J = 11.9$ and 5.4 , -CH_aH_b-N), 3.67 (1H, ddd, $J = 9.1$, 6.0 and 5.4 , >CH-N), 6.28 (1H, br d, $J = 6.0$, -NH-CO). ¹³C NMR (150.9 MHz, DMSO): 28.33 ((CH₃)₃), 40.65 (-CH₂-N), 51.08 (>CH-N), 78.42 (O-C(CH₃)₃), 155.35 (N-CO-O), 171.18 (-COOH). IR (KBr) ν_{\max} cm^{-1} 3348 m (NH); 1714 vs (C=O) acid; 1684 vs (C=O) carbamate; 1530 m (amide II); 1624 s (NH₂); 2978 m, 2932 m, 1366 m (CH₃). HRMS (ESI) calc for $\text{C}_{18}\text{H}_{15}\text{O}_4\text{N}_2$ [M-H]⁺ 203.10373, found: 203.10366.

Ac-β-azido-Ala-Val-Phe-CONH₂ **43**

Yield 115 mg (69%). Lyophilizate. $[\alpha]_D^{20} = -19.3$ (c = 0.218; DMSO). ¹H NMR (600 MHz, DMSO): 0.74 (3H, d, $J = 6.8$, CH₃), 0.75 (3H, d, $J = 6.8$, CH₃), 1.87 (3H, s, CH₃-CO), 1.92 (1H, m, >CH-), 2.78 (1H, dd, $J = 13.8$ and 9.1 , -CH_aH_b-), 2.99 (1H, dd, $J = 13.8$ and 5.2 , -CH_aH_b-), 3.38 (1H, dd, $J = 12.6$ and 8.0 , -CH_aH_b-N₃), 3.46 (1H, dd, $J = 12.6$ and 4.7 ,

–CHa**Hb**–N₃), 4.09 (1H, dd, $J = 8.6$ and 6.6 , >CH–N), 4.47 (1H, ddd, $J = 9.1$, 8.3 and 5.2 , >CH–N), 4.58 (1H, td, $J = 8.2$, 8.0 and 4.7 , >CH–N), 7.06 and 7.32 (2H, 2x br d, $J = 2.0$, CONH₂), 7.17 (1H, m, Ar-H), 7.22 (2H, m, Ar-H), 7.24 (2H, m, Ar-H), 7.90 (1H, d, $J = 8.3$, –NH–CO), 7.93 (1H, d, $J = 8.6$, –NH–CO), 8.31 (1H, d, $J = 8.2$, –NH–CO). ¹³C NMR (150.9 MHz, DMSO): 18.10 (CH₃), 19.30 (CH₃), 22.67 (CH₃–CO), 30.64 (>CH–), 37.82 (–CH₂–), 51.75 (–CH₂–N₃), 52.38 (>CH–N), 53.70 (>CH–N), 58.15 (>CH–N), 126.42, 128.23(2) and 129.31(2) (5x Ar =CH–), 137.95 (Ar >C=), 169.22 (N–CO–), 169.86 (N–CO–), 170.44 (N–CO–), 172.89 (–CONH₂). IR (KBr) ν_{\max} cm^{–1} 3423 s, 3292 s (NH); 2108 m (N₃); 1666 vs, 1642 vs (C=O); 1534 m (amide II); 3066 w, 3032 w, 1455 w, 1154 w, 700 w (ring); 2966 w, 2875 w (CH₃). HRMS (ESI) calc for C₁₉H₂₇O₄N₇Na [M+Na]⁺ 440.20167, found: 440.20177.

Ac-β-azido-Ala-Phe-Phe-CONH₂ **44**

Yield 141 mg (76%). Lyophilizate. $[\alpha]_D^{20} = -9.7$ ($c = 0.185$; DMSO). ¹H NMR (600 MHz, DMSO): 1.83 (3H, s, CH₃–CO), 2.74 (1H, dd, $J = 14.0$ and 9.2 , –CHa**Hb**–), 2.95 (1H, dd, $J = 14.0$ and 4.7 , –CHa**Hb**–), 2.82 (1H, dd, $J = 13.8$ and 8.7 , –CHa**Hb**–), 3.01 (1H, dd, $J = 13.8$ and 5.2 , –CHa**Hb**–), 3.31 (1H, dd, $J = 12.6$ and 8.1 , –CHa**Hb**–N₃), 3.42 (1H, dd, $J = 12.6$ and 4.6 , –CHa**Hb**–N₃), 4.43 (1H, ddd, $J = 8.7$, 8.3 and 5.2 , >CH–N), 4.44 (1H, ddd, $J = 9.2$, 8.1 and 4.7 , >CH–N), 4.48 (1H, ddd, $J = 8.4$, 8.1 and 4.6 , >CH–N), 7.10 and 7.28 (2H, 2x br d, $J = 2.0$, CONH₂), 7.16 – 7.27 (10H, m, 2x C₆H₅), 8.02 (1H, d, $J = 8.3$, –NH–CO), 8.15 (1H, d, $J = 8.1$, –NH–CO), 8.20 (1H, d, $J = 8.4$, –NH–CO). ¹³C NMR (150.9 MHz, DMSO): 22.68 (CH₃–CO), 37.46 (–CH₂–), 37.80 (–CH₂–), 51.70 (–CH₂–N₃), 52.23 (>CH–N), 53.88 (>CH–N), 54.31 (>CH–N), 126.48, 126.50, 128.25(2), 128.28(2) and 129.40(4) (10x Ar =CH–), 137.68 and 137.92 (2x Ar >C=), 169.07 (N–CO–), 169.79 (N–CO–), 170.52 (N–CO–), 172.77 (–CONH₂). IR (KBr) ν_{\max} cm^{–1} 3406 s, 3287 s (NH); 2107 m (N₃); 1666 vs, 1643 vs (C=O); 1538 m (amide II); 3064 w, 3031 w, 1498 w, 1454 w, 746 w, 700 w (ring). HRMS (ESI) calc for C₂₃H₂₆O₄N₇ [M+1]⁺ 464.20518, found: 464.20460.

Ac-γ-azido-Dab-Phe-Phe-CONH₂ **45**

Yield 153 mg (80%). Lyophilizate. $[\alpha]_D^{20} = -25.4$ ($c = 0.358$; DMSO). ¹H NMR (500 MHz, DMSO): 1.81 (3H, s, CH₃–CO), 1.63 and 1.75 (2H, 2x m, –CH₂–), 2.75 (1H, dd, $J = 13.9$ and 9.4 , –CHa**Hb**–), 2.95 (1H, dd, $J = 13.9$ and 4.6 , –CHa**Hb**–), 2.82 (1H, dd, $J = 13.9$ and 8.8 , –CHa**Hb**–), 3.01 (1H, dd, $J = 13.9$ and 5.1 , –CHa**Hb**–), 3.20 (2H, m, –CH₂–N₃), 4.24 (1H,

ddd, $J = 8.8, 7.8$ and 5.1 , $>CH-N$), 4.42 (1H, ddd, $J = 8.8, 8.1$ and 5.1 , $>CH-N$), 4.44 (1H, ddd, $J = 9.4, 8.0$ and 4.6 , $>CH-N$), 7.09 and 7.26 (2H, 2x d, $J = 1.9$, CONH₂), $7.16 - 7.27$ (10H, m, 2x C₆H₅), 7.98 (1H, d, $J = 8.0$, -NH-CO), 8.02 (1H, d, $J = 8.1$, -NH-CO), 8.03 (1H, d, $J = 7.8$, -NH-CO). ¹³C NMR (125.7 MHz, DMSO): 22.67 (CH₃-CO), 31.30 (-CH₂-), 37.38 (-CH₂-), 37.75 (-CH₂-), 47.64 (-CH₂-N₃), 50.31 ($>CH-N$), 53.96 ($>CH-N$), 54.09 ($>CH-N$), $126.45(2)$, $128.21(2)$, $128.27(2)$, $129.36(2)$ and $129.39(2)$ (10x Ar =CH-), 137.81 and 137.98 (2x Ar $>C=$), 169.68 (N-CO-), 170.75 (N-CO-), 171.07 (N-CO-), 172.80 (-CONH₂). IR (KBr) ν_{max} cm⁻¹ 3419 s, 3293 s (NH); 2104 m (N₃); 1665 s, 1643 vs (C=O); 1530 m (amide II); 3064 w, 3031 w, 1498 w, 1455 w, 746 w, 701 w (ring); 2930 w, (CH₂). HRMS (ESI) calc for C₂₄H₂₉O₄N₇Na [M+Na]⁺ 502.21732 , found: 502.21746 .

Ac- δ -azido-Orn-Phe-Phe-CONH₂ **46**

Yield 152 mg (77%). Lyophilizate. $[\alpha]_D^{20} = -23.6$ ($c = 0.233$; DMSO). ¹H NMR (600 MHz, DMSO): 1.81 (3H, s, CH₃-CO), 1.41 (2H, m, -CH₂-), 1.44 and 1.55 (2H, 2x m, -CH₂-), 2.75 (1H, dd, $J = 14.0$ and 9.4 , -CHaHb-), 2.96 (1H, dd, $J = 14.0$ and 4.6 , -CHaHb-), 2.83 (1H, dd, $J = 13.8$ and 8.7 , -CHaHb-), 3.00 (1H, dd, $J = 13.8$ and 5.2 , -CHaHb-), 3.24 (2H, m, -CH₂-N₃), 4.19 (1H, td, $J = 7.8, 7.8$ and 5.5 , $>CH-N$), 4.42 (1H, ddd, $J = 8.7, 8.3$ and 5.2 , $>CH-N$), 4.45 (1H, ddd, $J = 9.4, 8.1$ and 4.6 , $>CH-N$), 7.09 and 7.24 (2H, 2x d, $J = 1.9$, CONH₂), $7.16 - 7.26$ (10H, m, 2x C₆H₅), 7.977 (1H, d, $J = 7.8$, -NH-CO), 7.982 (1H, d, $J = 8.1$, -NH-CO), 7.989 (1H, d, $J = 8.3$, -NH-CO). ¹³C NMR (150.9 MHz, DMSO): 22.67 (CH₃-CO), 24.88 (-CH₂-), 29.30 (-CH₂-), 37.38 (-CH₂-), 37.74 (-CH₂-), 50.60 (-CH₂-N₃), 52.26 ($>CH-N$), 53.97 ($>CH-N$), 54.08 ($>CH-N$), 126.44 , 126.45 , $128.21(2)$, $128.27(2)$ and $129.37(4)$ (10x Ar =CH-), 137.86 and 137.97 (2x Ar $>C=$), 169.62 (N-CO-), 170.78 (N-CO-), 171.53 (N-CO-), 172.80 (-CONH₂). IR (KBr) ν_{max} cm⁻¹ 3397 m, 3283 m (NH); 2102 m (N₃); 1662 s, 1642 vs (C=O); 1543 m (amide II); 3065 w, 3031 w, 1498 w, 1455 w, 745 w, 700 w (ring); 2929 w, 2860 w (CH₂). HRMS (ESI) calc for C₂₅H₃₁O₄N₇Na [M+Na]⁺ 516.23297 , found: 516.23296 .

Ac- ω -azido-Lys-Phe-Phe-CONH₂ **47**

Yield 145 mg (71%). Lyophilizate. $[\alpha]_D^{20} = -24.5$ ($c = 0.326$; DMSO). ¹H NMR (600 MHz, DMSO): 1.20 (2H, m, -CH₂-), 1.40 and 1.50 (2H, 2x m, -CH₂-), 1.45 (2H, m, -CH₂-), 1.81 (3H, s, CH₃-CO), 2.75 (1H, dd, $J = 14.0$ and 9.5 , -CHaHb-), 2.95 (1H, dd, $J = 14.0$ and 4.7 ,

-CHa**Hb**-), 2.83 (1H, dd, $J = 13.8$ and 8.6 , -CHa**Hb**-), 3.00 (1H, dd, $J = 13.8$ and 5.1 , -CHa**Hb**-), 3.24 (2H, m, -CH₂-N₃), 4.16 (1H, ddd, $J = 8.6$, 8.0 and 5.3 , >CH-N), 4.42 (1H, ddd, $J = 8.6$, 8.0 and 5.1 , >CH-N), 4.44 (1H, ddd, $J = 9.5$, 8.1 and 4.7 , >CH-N), 7.10 and 7.24 (2H, 2x d, $J = 2.0$, CONH₂), 7.16 – 7.26 (10H, m, 2x C₆H₅), 7.94 (2H, 2x d, $J = 8.0$, 2x -NH-CO), 7.96 (1H, d, $J = 8.1$, -NH-CO). ¹³C NMR (150.9 MHz, DMSO): 22.65 (-CH₂-), 22.66 (CH₃-CO), 28.11 (-CH₂-), 31.54 (-CH₂-), 37.34 (-CH₂-), 37.75 (-CH₂-), 50.66 (-CH₂-N₃), 52.56 (>CH-N), 53.92 (>CH-N), 54.07 (>CH-N), 126.40, 126.43, 128.19(2), 128.25(2), 129.34(2) and 129.36(2) (10x Ar =CH-), 137.88 and 137.98 (2x Ar >C=), 169.55 (N-CO-), 170.77 (N-CO-), 171.81 (N-CO-), 172.74 (-CONH₂). IR (KBr) ν_{\max} cm⁻¹ 3291 m (NH); 2099 m (N₃); 1665 s, 1640 vs (C=O); 1539 m (amide II); 3065 w, 3031 w, 1498 w, 1455 w, 745 w, 701 w (ring); 2934 w, 2861 w (CH₂). HRMS (ESI) calc for C₂₆H₃₃O₄N₇Na [M+Na]⁺ 530,24862 found: 530.24861.

7.3 Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain

Background:

Insulin, IGF-1 and IGF-2 are structurally and functionally similar protein hormones. Their evolution originated from a common ancestor and has resulted in different yet overlapping biological functions. Insulin is a key regulator of basal metabolism, while IGF-1/2 are major growth factors. Insulin and IGFs cross-bind with different affinities to insulin receptor isoforms A and B (IR-A and IR-B) and IGF-1 receptor (IGF-1R), which mediate their specific biological functions. Identification of structural determinants in IGFs and insulin that trigger their specific signalling pathways is crucial for rational design of receptor specific analogues with potential therapeutic applications. The C domains of IGF-2 and IGF-1 are important structural determinants in which IGFs differ from insulin and each other.

Summary:

In this project, we developed a straightforward protocol for recombinant production of IGF-2 analogues and we used this new methodology for preparation of six IGF-2 analogues with IGF-1-like mutations in C domain. The analogues were tested for their binding features towards IR-A and IGF-1R. All of them exhibited significantly reduced affinity towards IR-A. Moreover, one of the analogues had also higher binding affinity for IGF-1R due to a synergistic effect of the Pro39-Gln40 insertion and Ser29Asn point mutation. Consequently this analogue had almost a 10-fold higher IGF-1R/IR-A binding specificity in comparison with native IGF-2. Moreover, we provided NMR structural characterization of selected IGF-2 analogues that helped to explain the binding behaviour of analogues by important conformational rearrangement of their C-domains. The development of the methodology for recombinant production of IGF-2 analogues made in this work was very important because later on we used the same methodology for the production of our first recombinant IGF-1 analogues.

My contribution:

I participated in the development of the methodology for the recombinant production of IGF-2 and analogues and I personally prepared two IGF-2 analogues (six in total) by this

methodology. I also contributed to the performing of the binding tests with the IGF-2 analogues.

Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain^{*[S]}

Received for publication, June 2, 2016, and in revised form, July 29, 2016. Published, JBC Papers in Press, August 10, 2016, DOI 10.1074/jbc.M116.741041

Rozálie Hexnerová^{‡§1}, Květoslava Křížková^{‡§1}, Milan Fábry^{‡¶}, Irena Siegllová[‡], Kateřina Kedrová^{‡§5}, Michaela Collinsová[‡], Pavlína Ullrichová^{||}, Pavel Srb[‡], Christopher Williams^{**}, Matthew P. Crump^{**}, Zdeněk Tošner[§], Jiří Jiráček[‡], Václav Veverka^{‡,2}, and Lenka Žáková^{‡,3}

From the [‡]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., Flemingovo nám 2, 166 10 Prague 6, Czech Republic, [§]Faculty of Science, Charles University in Prague, Albertov 6, Prague 128 43, Czech Republic, ^{||}Department of Analytical Chemistry, University of Chemistry and Technology, Technická 5, 166 28 Prague 6, Czech Republic, [¶]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, v.v.i., Vídeňská 1083, 142 20 Prague 4, Czech Republic, and ^{**}Department of Organic and Biological Chemistry, School of Chemistry, Cantock's Close, University of Bristol, Bristol BS8 1TS, United Kingdom

Insulin and insulin-like growth factors I and II are closely related protein hormones. Their distinct evolution has resulted in different yet overlapping biological functions with insulin becoming a key regulator of metabolism, whereas insulin-like growth factors (IGF)-I/II are major growth factors. Insulin and IGFs cross-bind with different affinities to closely related insulin receptor isoforms A and B (IR-A and IR-B) and insulin-like growth factor type I receptor (IGF-1R). Identification of structural determinants in IGFs and insulin that trigger their specific signaling pathways is of increasing importance in designing receptor-specific analogs with potential therapeutic applications. Here, we developed a straightforward protocol for production of recombinant IGF-II and prepared six IGF-II analogs with IGF-I-like mutations. All modified molecules exhibit significantly reduced affinity toward IR-A, particularly the analogs with a Pro-Gln insertion in the C-domain. Moreover, one of the analogs has enhanced binding affinity for IGF-1R due to a synergistic effect of the Pro-Gln insertion and S29N point mutation. Consequently, this analog has almost a 10-fold higher IGF-1R/IR-A binding specificity in comparison with native IGF-II. The established IGF-II purification protocol allowed for cost-effective isotope labeling required for a detailed NMR structural

characterization of IGF-II analogs that revealed a link between the altered binding behavior of selected analogs and conformational rearrangement of their C-domains.

The insulin-insulin-like growth factor (IGF)⁴ axis is a complex signaling pathway mediated by a group of three sequentially and structurally homologous peptide hormones, their membrane receptors, and several circulating IGF-binding proteins. Insulin and IGF-I and -II are all capable of higher or lower affinity binding toward the transmembrane tyrosine kinase receptors insulin receptor isoform A (IR-A), insulin receptor isoform B (IR-B), and insulin-like growth factor type I receptor (IGF-1R) (1, 2). All three receptors also share a high degree of homology, which is manifested by overlapping biological responses upon ligand binding (3–5). Binding of insulin and IGFs to the receptors triggers two major signaling pathways via autophosphorylation of tyrosines within their intracellular tyrosine kinase domains. The first, usually referred to as a phosphoinositide 3-kinase (PI3K)/Akt pathway, is key for the metabolic effects of ligand binding such as a decrease in plasma glucose levels (6). The second signaling pathway, referred to as Ras/ERK, involves activation of the Ras/Raf/MAPK/ERK1/2 cascade, which mediates proliferative effects through gene transcription regulation (7). Whereas insulin signals mainly via both IR isoforms (8), IGF-I and IGF-II promote the mitogenic signaling through IGF-1R (9, 10), and similar mitogenic stimulation results from IGF-II binding to IR-A (11).

Both IGFs are essential for embryonic development and are present in serum at nanomolar concentrations in adults (12) with IGF-II levels being 3-fold higher than IGF-I levels (13). Whereas the role of IGF-II in tumor development is well doc-

* This work was supported by Czech Science Foundation Grant 15-19018S, Medical Research Council Grant MR/K000179/1, Ministry of Education of the Czech Republic Programs "NAVRAT" LK11205 and "NPU I" LO1304, Charles University Grant Agency Grant 227020, Specific University Research (Ministry of Education of the Czech Republic Grant 20/2013, A1_FCHI_2014_003), and Research Project of the Academy of Sciences of the Czech Republic RVO:61388963. The authors declare that they have no conflicts of interest with the contents of this article.

§ Author's Choice—Final version free via Creative Commons CC-BY license.

[S] This article contains supplemental Figs. S1–S8 and Table S1.

The atomic coordinates and structure factors (codes 5L3L, 5L3M, and 5L3N) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

The assigned chemical shifts have been deposited into the BioMagResBank under accession numbers 34000, 34001, and 34002.

¹ Joint first authors.

² To whom correspondence may be addressed. Tel.: 420-220-183-135; E-mail: vaclav.veverka@uochb.cas.cz.

³ To whom correspondence may be addressed: Inst. of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., Flemingovo nám 2, 166 10 Prague 6, Czech Republic. Tel.: 420-220-183-441; E-mail: zakova@uochb.cas.cz.

⁴ The abbreviations used are: IGF, insulin-like growth factor; IR, insulin receptor; IR-A, insulin receptor isoform A; IR-B, insulin receptor isoform B; IGF-1R, insulin-like growth factor type I receptor; IGF-2R, insulin-like growth factor type II receptor; L1, leucine-rich repeat region; α -CT, C-terminal helix; GB1, immunoglobulin binding domain B1 of streptococcal Protein-G; TEV, tobacco etch virus; RP-HPLC, reversed phase HPLC; HSQC, heteronuclear single quantum coherence; D11, Domain 11; HMQC, heteronuclear multiple quantum coherence.

Receptor Specificity of IGF-II Analogs

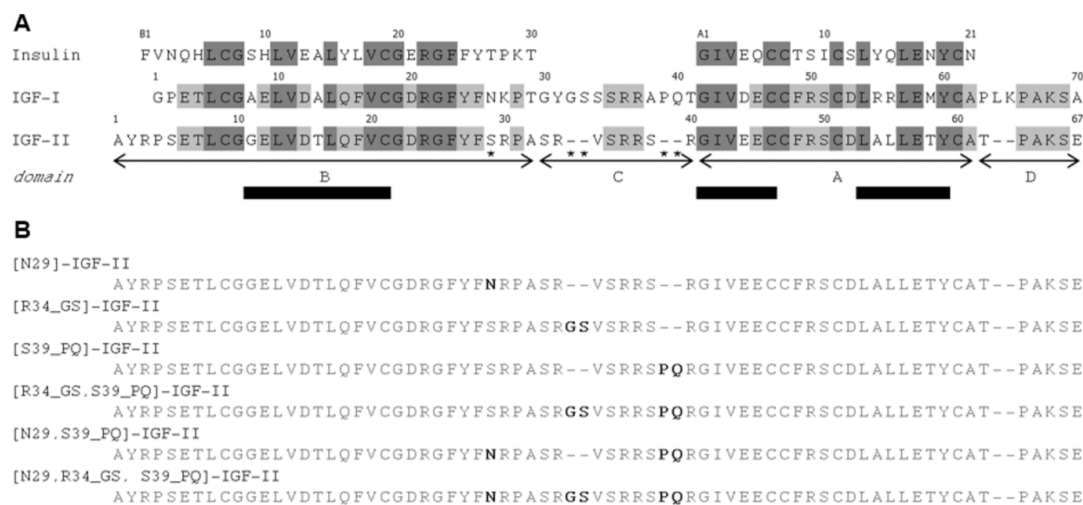


FIGURE 1. A, the amino acid sequence alignment of insulin, IGF-I, and IGF-II. It illustrates their high primary structure homology with the conserved residues highlighted in *dark gray* and the residues conserved between IGF-I and IGF-II in *light gray*. The organization of IGF-I and IGF-II into B-, C-, A-, and D-domains is shown below the sequences; domains A and B correspond to insulin A and B chains. The positions of conserved α -helices are shown as *bars* above the sequences. IGF-II residues mutated in this study are labeled with an *asterisk*. B, the amino acid sequence of the six prepared IGF-II analogs with highlighted mutations.

umented (14), its physiological role remains unclear. It is known that IGF-II is important for fetal development and placental function (15, 16), and several animal studies indicate an important role for IGF-II in memory enhancement (17–19). The availability of IGF ligands for signaling is modulated by a family of high affinity IGF-binding proteins 1–6 (20, 21) and insulin-like growth factor type II receptor (IGF-2R) (22). The equilibrium of individual components and the appropriate function of the entire insulin-IGF system are essential for biological responses such as regulation of basal metabolism, cellular growth, proliferation, survival, and migration (23).

IGF-I and IGF-II are single chain peptides composed of 70 and 67 amino acids, respectively. Mature IGFs consist of four domains: B, C, A, and D in order from the N terminus. IGF-I and -II share over 60% sequence identity, mostly in the B- and A-domains that correspond to the B and A chains in insulin (Fig. 1). The 3D structure of IGF-I was obtained by both NMR and x-ray (24–34), whereas the structure of IGF-II has been determined only by NMR (35, 36). Together with insulin, these hormones share the insulin-like conformation consisting of three highly conserved α -helices (Fig. 1) further stabilized by three characteristic disulfide bonds (28, 36, 37).

IR-A, IR-B, and IGF-1R are homodimeric, and each monomer consists of an extracellular subunit (α) and transmembrane subunit (β) that are linked via four disulfide bonds into a functional β - α - β homodimer (38–40). The alternative splicing of IR exon 11 generates a 12-amino acid sequence in the C terminus of the α -subunit or IR-B that is absent in IR-A (41–43). Each monomer contains two insulin/IGF binding sites termed the primary (1) and second (2) site on one monomer and 1' and 2' on the partner. The primary binding site is formed from a leucine-rich repeat region (L1) and C-terminal helix (α -CT) region that combine with the second site of the partner monomer (2') to form the complete binding pocket. The two sites

(1-2') bind a single molecule of insulin/IGF, triggering structural rearrangements and negative cooperativity for binding at the 1'-2 site (44–46). The mechanisms of insulin or IGF binding to their cognate receptors were originally proposed on the basis of extensive mutagenesis studies only (47, 48). More recently, however, several reports based on the crystal structures of the insulin-IR complexes (49, 50), "activated" insulin analogs (51–53), and the first bound structure of IGF-I through complexation with a IR/IGF-1R hybrid construct (54) have revealed the binding mode of the hormones at the receptor site 1 represented by the L1 subunit and α -CT segment. However, details of the precise arrangement of the C-domain of bound IGF-I are currently unknown, but structural rearrangement of this region in conjunction with the α -CT region has been proposed to be necessary to prevent unfavorable steric clashes. Moreover, the C-domain is a region with major differences between IGFs, both in the amino acid composition and length (Fig. 1), probably being a key determinant of receptor binding specificity.

Both insulin and IGF-I have been extensively studied through the preparation and functional analysis of numerous analogs (for extensive reviews, see Refs. 46, 48, and 55), whereas the structure-function of IGF-II is less developed (15, 56–62). To gain greater insight into the structural basis of IGF-II binding specificity to IR-A and IGF-1R, we generated a series of mutants containing amino acid substitutions within the B- and C-domains of IGF-II. These were designed to make IGF-II more IGF-I-like (Fig. 1) and were tested through binding affinities to their cognate receptors. This was enabled by the development of a new, efficient, and cost-effective protocol for recombinant production of IGF-II analogs in sufficient quantities for structural characterization by NMR. Our data revealed that the newly prepared IGF-II analogs display conserved or slightly increased IGF-1R affinities

Receptor Specificity of IGF-II Analogs

but markedly reduced IR-A affinities, which correlates with the specific conformational changes in the structurally elusive C-domain of IGF-II.

Results

Recombinant Production of IGF-II—A prerequisite for this study was the efficient production of correctly folded IGF-II, which would serve as a reference molecule as well as a platform for the design and production of new IGF-II analogs. This was achieved by recombinant IGF-II expression in *Escherichia coli* as a fusion with an N-terminal and cleavable His₆-tagged GB1 protein (immunoglobulin binding domain B1 of streptococcal Protein-G) (63, 64). This technique provided high yields (0.8–1.8 mg liter⁻¹ of culture) of IGF-II analogs with only a single additional glycine residue cloning artifact at the N terminus. The fusion protein was successfully expressed in *E. coli* and purified using immobilized metal ion affinity chromatography (supplemental Fig. S1). Two major peaks were observed; the first eluted at lower concentrations of imidazole (110–160 mM; fractions 1–2 in supplemental Fig. S1) and consisted of folded and misfolded monomeric IGF-II with slightly different migration of two bands observable in non-reducing SDS-PAGE (supplemental Fig. S1). The second peak eluted at higher concentrations of imidazole and consisted of multimeric forms (310–480 mM; fractions 4–5 in supplemental Fig. S1). Both monomeric and multimeric fusion proteins were subsequently cleaved using TEV protease under redox conditions of oxidized and reduced glutathione. Interestingly, the moderate reducing environment triggered disulfide bond reshuffling that resulted in liberation of monomeric IGF-II from multimeric aggregates. Following cleavage, IGF-II was separated from the His₆-tagged GB1 and TEV by immobilized metal ion affinity chromatography. RP-HPLC of this crude IGF-II product consisted of one major peak and two to four minor peaks (supplemental Fig. S1). The retention time of the major protein peak was nearly identical to that observed for native human IGF-II, and the correct molecular weight of the recombinantly produced purified IGF-II protein with formed disulfide bonds was confirmed by high resolution mass spectrometry. Both forms, monomeric and multimeric, yielded the desired product of correct mass and were combined after the correct protein fold was confirmed by 1D ¹H NMR (supplemental Fig. S2) and ¹H-¹⁵N HSQC that was highly similar to the previously published data (65).

In total, six IGF-II analogs were designed to determine the effects of IGF-I motif incorporation into IGF-II. The modifications were as follows: (i) a point mutation at position Ser²⁹ for Asn ([N29]IGF-II), (ii) an insertion of Gly-Ser after Arg³⁴ ([R34_GS]IGF-II), (iii) an insertion of Pro-Gln after Ser³⁹ ([S39_PQ]IGF-II), (iv) a combination of both insertions ([R34_GS,S39_PQ]IGF-II), (v) a combination of S29N mutation with Pro-Gln insertion ([N29,S39_PQ]IGF-II), and (vi) a combination of S29N mutation with both insertions ([N29,R34_GS,S39_PQ]IGF-II). All analogs gave comparable RP-HPLC elution profiles (data not shown) with that of IGF-II (supplemental Fig. S1) with one major product and several minor peaks. The characterization of minor by-products was prevented by their relatively low yields.

The structural integrity of the six analogs was confirmed using ¹H NMR and far-UV circular dichroism as illustrated in supplemental Figs. S2 and S3. The CD spectra obtained for prepared analogs are similar to the broadly α -helical secondary structure profile obtained for non-modified IGF-II. The presence of the expected tertiary structure was further confirmed by 1D ¹H (supplemental Fig. S2) NMR spectra, and each analog compared well with the native IGF-II profile.

Receptor Binding—The binding affinities of the IGF-II analogs toward human IR-A and IGF-1R together with binding affinities of selected analogs to IR-B are summarized in Table 1 and Fig. 2. The corresponding binding curves are shown in supplemental Figs. S4–S6.

IR-A Binding Affinities—All modifications led to a significantly impaired IR-A binding, ranging from 4.2 to 1.1% of the insulin affinity when compared with IGF-II (7.9%). The [N29]IGF-II B-domain mutant gave a 2-fold reduction in IR-A affinity, whereas the analogs with C-domain insertions exhibited stronger negative effects. [R34_GS]IGF-II showed an almost 3-fold reduction in binding (2.8%), whereas [S39_PQ]IGF-II showed an 8-fold reduction. All of the analogs bearing the Pro-Gln motif were significantly less active (1.1–1.8%), and further combinations did not appear to have any additive effect.

IGF-1R Binding Affinities—An insertion of IGF-I-like features, S29N, Gly-Ser, Pro-Gln alone, or a combination of Gly-Ser and Pro-Gln, within the IGF-II molecule led rather unexpectedly to a moderate decrease of binding potency toward IGF-1R (Table 1 and Fig. 2). However, the Pro-Gln insertion combined with the S29N mutation resulted in an increase in binding potency to that of 18.8% to IGF-I in comparison with IGF-II (10.9%). In contrast, this effect was negated when the S29N mutation was combined with both insertions.

IR-B Binding Affinities—Both reference molecules, commercial IGF-II and our recombinant IGF-II, show similar binding potency for IR-B compared with IGF-I (1.9 and 1.5% of human insulin, respectively; ~40 nM; Table 1). The IR-B binding affinity of [N29]IGF-II dropped to almost one-third of the potency obtained for IGF-II (0.6%; 108 nM).

Structural Characterization of IGF-II Analogs by NMR Spectroscopy—We selected two IGF-II analogs with the most pronounced impact on receptor binding [S39_PQ]IGF-II (with lowest IR-A and IGF-1R binding) and [N29,S39_PQ]IGF-II (with decreased IR-A and enhanced IGF-1R binding) (Table 1 and Fig. 2) for NMR structural characterization to understand the molecular basis of Pro-Gln and S29N modifications.

Undesirable dynamic and aggregation behavior of IGF-II severely affects the quality of NMR spectra of this protein and would prevent the accurate structural determination required for a detailed comparison between these analogs. Previously, it has been shown that upon binding to an engineered high affinity Domain 11 (D11) of the IGF-2R the spectral properties of IGF-II improve dramatically (65). The fact that the IGF-II modifications reported here are distributed on the opposing face to the D11 binding site allowed this system to be utilized for structural studies of the B- and C-domains. As expected, the binding of either ¹⁵N- or ¹³C/¹⁵N-labeled IGF-II analogs to unlabeled D11 led to a significant line narrowing of the NMR signals as illustrated in supplemental Fig. S7 despite the more than a

TABLE 1

The receptor binding affinities of hormones and IGF-II analogs reported in this work

The values of K_d and relative binding affinities of human insulin, IGF-I, IGF-II, and IGF-II analogs were determined for human IR-A in membranes of human IM-9 lymphocytes and for human IR-B and human IGF-IR in membranes of mouse fibroblasts. Relative receptor binding affinity is defined as $(K_d/\text{of human insulin or IGF-II analog}) \times 100$. ND is not determined.

Analog	$K_d \pm$ S.E. (n) for human IR-A in IM-9 lymphocytes	Relative binding affinity for human IR-A	$K_d \pm$ S.E. (n) for human IGF-IR in mouse fibroblasts	Relative binding affinity for human IGF-IR	$K_d \pm$ S.E. (n) for human IR-B in mouse fibroblasts	Relative binding affinity for human IR-B
Commercial human insulin	0.43 ± 0.02 (5)	100 ± 5	292 ± 31 (3) ^a	0.08 ± 0.01	0.67 ± 0.17 (5) ^b	100 ± 18
	0.24 ± 0.02 (5)	100 ± 8	0.24 ± 0.05 (5) ^a	100 ± 21	0.67 ± 0.08 (4) ^a	100 ± 12
	23.8 ± 6.6 (3) ^b	1.0 ± 0.3 ^c	0.25 ± 0.01 (4)	100 ± 4	224 ± 16 (4) ^b	0.3 ± 0.0 ^d
Commercial human IGF-I			0.25 ± 0.01 (4)	10.8 ± 3.3 ^e	35.5 ± 5.6 (4) ^b	1.9 ± 0.3 ^d
Commercial human IGF-II			2.32 ± 0.72 (3) ^b	10.9 ± 5.0 ^e	43.7 ± 5.3 (3)	1.5 ± 0.2 ^d
IGF-II			2.29 ± 1.04 (4)	5.3 ± 1.3 ^f	108 ± 16 (3)	0.6 ± 0.1 ^h
[N29]IGF-II			4.57 ± 1.09 (3)	5.8 ± 1.3 ^f	ND	ND
[R34_GS]IGF-II			4.13 ± 0.90 (3)	4.8 ± 1.1 ^g	ND	ND
[S39_PQ]IGF-II			5.00 ± 1.10 (3)	4.2 ± 1.6 ^g	ND	ND
[R34_GS,S39_PQ]IGF-II			5.68 ± 2.13 (3)	18.8 ± 5.1 ^e	ND	ND
[N29,S39_PQ]IGF-II			1.33 ± 0.36 (3)	7.8 ± 2.7 ^e	ND	ND
[N29,R34_GS,S39_PQ]IGF-II			3.19 ± 1.08 (3)		ND	ND

^a From Vikova *et al.* (87).

^b From Křížková *et al.* (2).

^c Relative to human insulin K_d value of 0.24 ± 0.02 ($n = 5$).

^d Relative to human insulin K_d value of 0.67 ± 0.12 ($n = 5$).

^e Relative to human IGF-I K_d value of 0.25 ± 0.01 ($n = 4$).

^f Relative to human insulin K_d value of 0.43 ± 0.02 ($n = 5$).

^g Relative to human IGF-I K_d value of 0.24 ± 0.05 ($n = 5$).

^h Relative to human insulin K_d value of 0.67 ± 0.08 ($n = 4$).

ⁱ This K_d value represents mean of two independent measurements ± range.

2-fold increase in the total molecular mass of the system. First, we determined the structure of the D11-bound unmodified IGF-II that was utilized in the structural analysis of IGF-II analogs. As expected, it is highly similar to the previously published structure (65) with some regions being more resolved, especially around the sites modified in the analogs, reflecting the substantially higher number of experimental restraints (1039 versus 764 unambiguous NOE restraints (supplemental Table S1 and Ref. 65)). Next, we verified that binding to D11 did not significantly affect the IGF-II C-domain and C-terminal portion of the B-domain by comparison of assigned 2D ¹H-¹⁵N HSQC spectra of free and D11-bound [S39_PQ]IGF-II (supplemental Fig. S8). Although significant chemical shift perturbations were observed over the A-domain and the first 75% of the B-domain, the regions containing the mutations showed very small or negligible chemical shift perturbations.

Both analogs, [S39_PQ]IGF-II and [N29,S39_PQ]IGF-II, preserved their overall structural organization with the three highly conserved α -helices further stabilized by three disulfide bonds. As expected, the D11 binding interface on the IGF-II analogs was not perturbed by the modifications, and structural changes were restricted to the modification sites (Fig. 3). In both analogs, the C-domain insertion led to a significant change in the conformational space sampled by this region of the protein compared with unmodified IGF-II with the main differences residing between residues 29 and 42. Detailed analysis (Fig. 4) revealed that the insertion of Pro-Gln after Ser³⁹ led to increased conformational freedom within the C-loops of both analogs that generated a rearrangement stabilized by several new packing interactions in the remote part of the C-domain. In the native IGF-II sequence, Tyr²⁷ points away from the C-loop and forms hydrophobic contacts with Ala⁶¹, whereas the C-loop is unrestrained by additional contacts to the other parts of IGF-II (Fig. 4A). By contrast, the aromatic ring of Tyr²⁷ forms contacts to the methyl group of Ala³² in [S39_PQ]IGF-II (Fig. 4B) and Arg³⁰ and Pro³¹ in [N29,S39_PQ]IGF-II (Fig. 4C). Arg³⁰ is no longer unrestrained in these analogs and interacts with the aromatic ring of Tyr⁶¹ (Tyr⁵⁹ in unmodified IGF-II) via a cation- π interaction. These new hydrophobic contacts lead to the formation of a better defined C-loop that bends around the bulky side chains of Tyr²⁷ and Tyr⁶¹ of both C-domain-modified analogs (Fig. 4, B and C). In comparison with unmodified IGF-II, the extended C-domain in both analogs is spatially constrained and bent toward the triad of aromatic residues at the C terminus of the B-domain (Phe²⁶, Tyr²⁷, and Phe²⁸). Ser²⁹ in IGF-II (Fig. 4A) is located at the hinge of the semiflexible loop with no significant contacts to neighboring residues. The Pro-Gln extension in [S39_PQ]IGF-II led to the repositioning of Ser²⁹ in close proximity to Tyr²⁷, although there are no observed NOE contacts between Ser²⁹ protons and Tyr²⁷ or surrounding residues. However, the hydroxyl proton from its side chain may be involved in hydrogen bonds, e.g. with the backbone carboxyl groups either from Pro³¹ (<2.8 Å in half of the structures), which is closer in the extended loop, or from Arg³² (<2.8 Å in a quarter of the structures) at the opposite side of the loop (Figs. 4B and 5). The modification of Ser²⁹ to Asn²⁹ in [N29,S39_PQ]IGF-II led to a loss of this hydrogen bond and a subtle conformational rearrangement of the C-loop backbone. In addition, the Asn²⁹ side chain is pointing out of the C-loop

Receptor Specificity of IGF-II Analogs

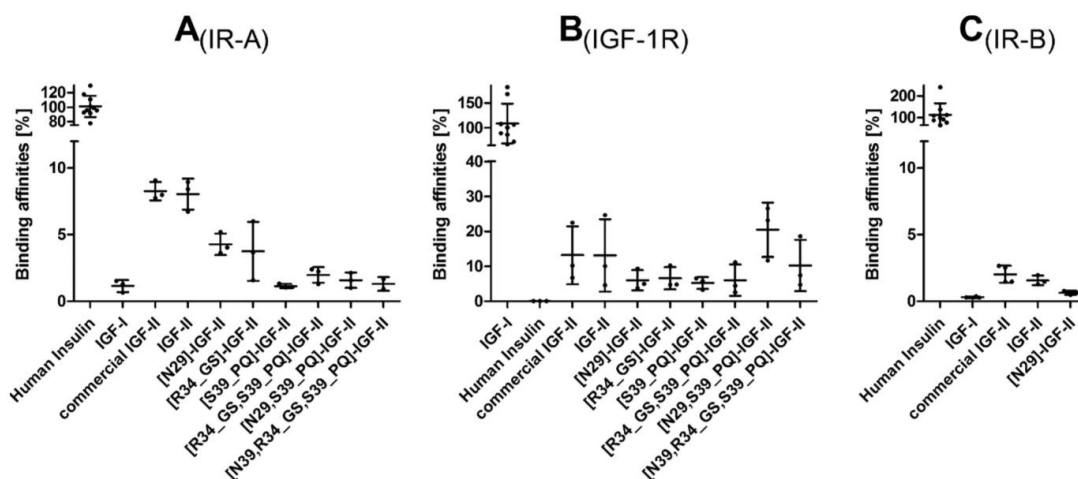


FIGURE 2. **Summary of receptor binding affinities.** Shown is a bar plot representation of relative binding affinities (from Table 1) of native hormones and IGF-II analogs prepared in this work for human IR-A (A), IGF-1R (B), and IR-B (C). Error bars represent S.D.

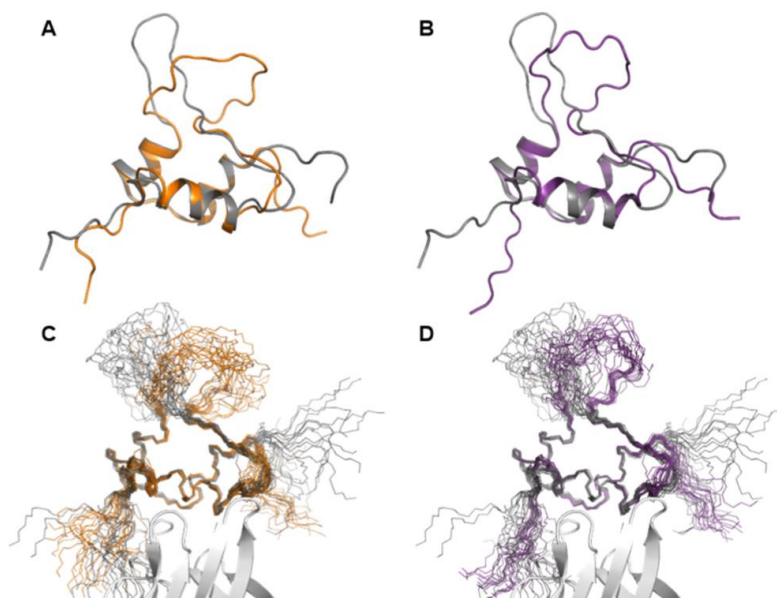


FIGURE 3. **Solution structures of [S39_PQ]IGF-II (orange) and [N29,S39_PQ]IGF-II (purple) compared with non-modified IGF-II (gray).** A and B show representative structures of the Domain 11-bound IGF-II analogs, and C and D show sets of 20 converged structures bound to D11 (white). The insertion of Pro-Gln in the C-domain after position 39 led to a significant structural rearrangement of the semiflexible loop.

and is fully solvent-exposed with NOE contacts between the NH_2 group from the Asn²⁹ side chain and H ^{β} 2 from Phe²⁸, perhaps further stabilizing the cluster of contacts between the C-domain and aromatic triad that in turn might stabilize the additional interactions seen between Tyr²⁷ and Arg³⁰/Pro³¹ (Fig. 4C) that were not observed for the [S39_PQ]IGF-II analog.

Discussion

IGF-II is capable of binding to both IR-A and IGF-1R with single digit nanomolar affinity ($K_d \sim 3$ nM; Table 1) and to IR-B

with lower affinity (~ 40 nM; Table 1). Although the binding affinities of the “parent” ligands, insulin and IGF-I, toward their cognate receptors are in the subnanomolar range (Table 1), IGF-II can still effectively signal through both IR-A and IGF-1R receptors or their hybrid forms *in vivo* (66, 67), which may trigger unfavorable biological responses. The knowledge about structural elements within these hormones responsible for differential binding specificity to each receptor could open a new path to the development of receptor-selective IGF and insulin analogs with potential medical applications. The analogs pre-

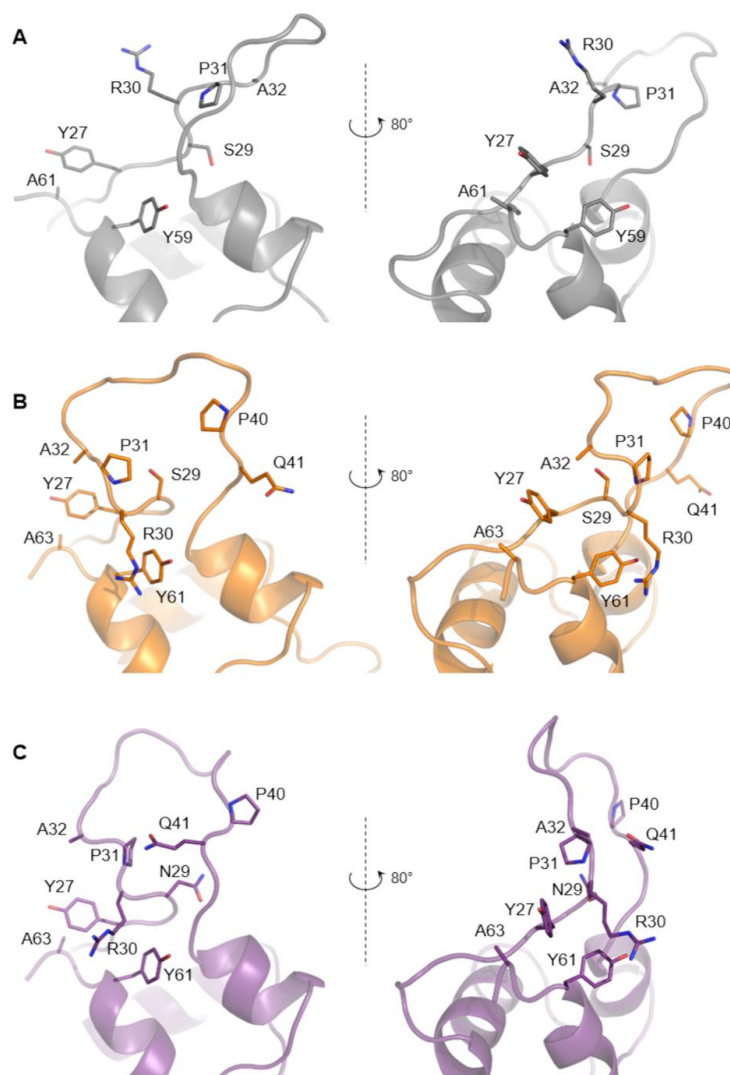


FIGURE 4. **Structural impact of the IGF-II modifications.** Non-modified IGF-II (A; gray) is compared with [S39_PQ]IGF-II (B; orange) and [N29,S39_PQ]IGF-II (C; purple), revealing different spatial orientation of highlighted residues. In particular, the rearrangement of the C-domain is driven by repositioning of Ala³² toward Tyr²⁷ and Arg³⁰ toward Tyr⁶¹ (Tyr⁵⁹ in non-modified IGF-II) supported by additional contacts within this area.

pared and structurally characterized in this work were designed to investigate the effects of introducing unique IGF-I motifs (*i.e.* Asn²⁶, Gly³⁰-Ser³¹, and Pro³⁵-Gln³⁶; Fig. 1) to IGF-II on receptor binding behavior. We hypothesized that such modifications may negatively affect the hormone's binding potency toward IR-A while enhancing the binding affinity for IGF-1R. Moreover, there are no reported analogs with the mutation of Ser²⁹ in IGF-II, and there are only a few studies regarding alterations in the C-domain (57, 59).

The development of an efficient protocol for IGF-II production was a key step in being able to reliably prepare the IGF-II analogs. The total chemical synthesis of IGF-II is extremely

difficult and time-consuming due to the length and unfavorable composition of the IGF-II sequence (68). The most frequently used recombinant approach, analogous to the production of IGF-I (69, 70), is based on preparation of a fusion comprising porcine growth hormone N-terminal residues 1–11 (plus N-terminal Met), a subtilisin-specific cleavage sequence (Val-Asn-Phe-Ala-His-Tyr ↓), and human IGF-II (71). However, specifically mutated subtilisin (H64A) used for the procedure is no longer commercially available. We therefore chose an alternative approach that includes an “on-column” refolding step of denatured IGF-II in a fusion with His₆-tagged GB1 protein (63, 64). Subsequent cleavage of the fusion protein in a redox envi-

Receptor Specificity of IGF-II Analogs

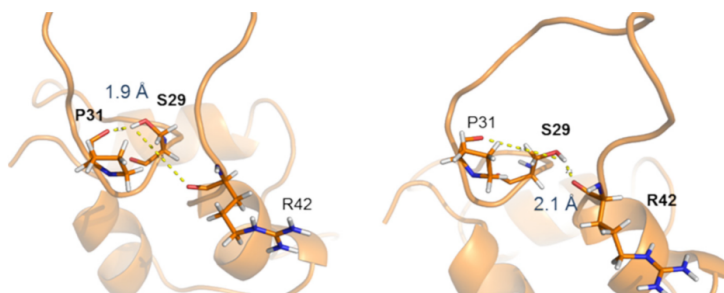


FIGURE 5. **The formation of stabilizing hydrogen bond in [S39_PQ]IGF-II.** The hydroxyl proton from the Ser²⁹ side chain is stabilizing the C-loop via a hydrogen bond to the backbone carboxyl groups either from Pro³¹ or Arg⁴².

ronment and RP-HPLC purification yields IGF-II with only a single additional glycine residue at the N terminus. This improves on the recently reported recombinant method that leaves three surplus N-terminal amino acids (glycine, alanine, and methionine) (65, 72) and therefore reduces uncertainty in interpreting structure and function of this protein in biological assays.

The binding affinities of recombinantly produced IGF-II toward IR-A, IR-B, and IGF-1R correlate with the values obtained for commercial IGF-II (Table 1 and Fig. 2). These comparable binding characteristics confirmed the correct disulfide pairing as misfolded IGFs do not bind to IGF-1R or IR-A with a measurable affinity (27, 73, 74). This method therefore leads to a rapid and cost-effective preparation of authentic IGF-II, providing us and others with an essential tool for studying IGF-II-related structure and function.

Our initial goal to reduce IR-A affinity of IGF-II was successful as all six IGF-II analogs showed reduced IR-A binding (Table 1) with four of these showing low affinity similar to IGF-I. The most significant was [S39_PQ]IGF-II with an ~7–8-fold reduction in affinity compared with IGF-II. Interestingly, this reduction was greater than the effect of swapping the entire IGF-II C-domain for IGF-I C-domain (3.7-fold) (57). Our data and data of others (57, 59) suggest that there are two main factors affecting the binding potency of IGF-II to IR-A. First, a longer C-domain may introduce structural restrictions during binding to the insulin receptor. This is in agreement with the finding that IGF-I analogs with a shorter C-domain exhibit enhanced binding affinity to IR-A (75). The second factor relates to specific C-domain amino acids (e.g. Pro³⁹ in IGF-I), which may affect the structure of the C-domain main chain and therefore binding to IR-A.

Although we only tested a single analog for binding to IR-B, the 2-fold reduction in binding observed for [N29]IGF-II suggests a similar sensitivity to changes in the C-domain (Table 1). Hence, we have not further pursued testing IR-B affinities of analogs and we focused on binding to IGF-1R.

The combination of the Pro-Gln insertion with S29N mutation in [N29,S39_PQ]IGF-II (Table 1 and Fig. 2) led to an analog exhibiting higher binding affinity to IGF-1R compared with native IGF-II. Our data suggest that the IGF-II specificity toward IGF-1R is determined by the amino acid composition of the C-domain rather than its length as demonstrated by the relatively lower binding affinity of the [R34_GS,S39_PQ]IGF-II

analog. The selected mutations do not completely recover IGF-I-like binding to IGF-1R and cannot counterbalance the absence of other important IGF-I determinants (e.g. IGF-I Tyr³¹ (76–78)). Nonetheless, the almost doubling in IGF-1R binding affinity of [N29,S39_PQ]IGF-II analog together with its markedly lowered affinity for IR-A resulted in almost 10-fold enhanced IGF-1R/IR-A binding specificity in comparison with IGF-II.

The comparison of D11-bound structures of IGF-II, [S39_PQ]IGF-II and [N29,S39_PQ]IGF-II, revealed that both analogs differ from IGF-II in the orientation and structuring of their C-loops (Figs. 3 and 4). The significant and similar displacement of the C-loops in both [S39_PQ]IGF-II and [N29,S39_PQ]IGF-II together with their more open C-loop conformations can be attributed to the effect of their PQ inserts. Moreover, the C-loop loops back to generate a turn stabilized by contacts between Tyr²⁷ and Ala³² and a hydrogen bond between Ser²⁹ and Pro³¹ or Arg⁴² in [S39_PQ]IGF-II (Figs. 4B and 5). The absence of this hydrogen bond due to the S29N mutation in [N29,S39_PQ]IGF-II might be compensated for by Pro³¹ packing against Tyr²⁷ (Fig. 4C). A comparable decrease in IR-A binding affinities of [S39_PQ]IGF-II and [N29,S39_PQ]IGF-II in comparison with IGF-II indicates it is caused mainly by their similarly altered C-loop structures rather than S29N mutation, which is well tolerated by IR-A.

In the crystal structure of human IGF-I (Protein Data Bank code 1GZR) (29), the Asn²⁶ side chain is solvent-exposed at the interface of the B- and C-domains with the Asn²⁶ presenting a potential polar hot spot. An equivalent Asn²⁹ in [N29,S39_PQ]IGF-II is in a similar position but is less exposed due to a partial overlap by the rearranged C-loop (Fig. 6A). Asn²⁶ is at the C terminus of the IGF-I B-domain, which is structurally altered upon binding to IGF-1R or IR (54) (Fig. 6, IGF-I receptor-bound structures in cyan). Analogous structural events are observed upon insulin binding to IR (50, 53), and it can be expected that receptor-driven activation of IGF-II is similar. In the Menting *et al.* (54) structure (Protein Data Bank code 4XSS), Asn²⁶ is the last IGF-I B-domain residue resolved in the complex with the hybrid IGF-IR/IR where it has been captured in the binding site formed from the IGF-IR α -CT and IR L1 domains (Fig. 6B). However, the structure of the complex did not reveal any specific contacts between IGF-I Asn²⁶ and IR L1 domain or IGF-IR α -CT. However, it cannot be ruled out that Ser²⁹ within the IGF-II molecule (or Asn²⁹ in

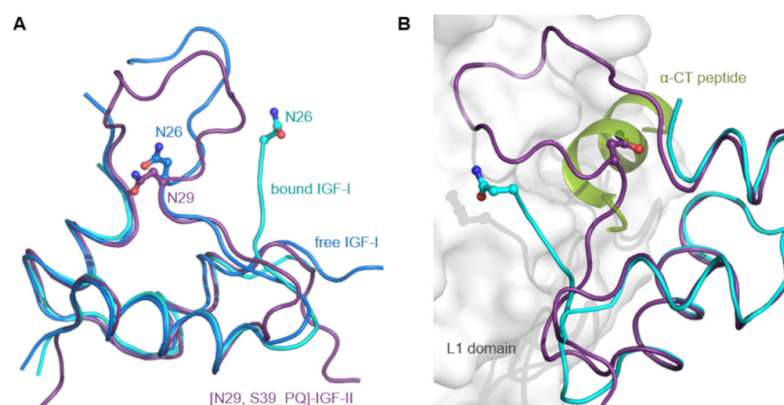


FIGURE 6. **A** superposition of free or hybrid IR/IGF-1R fragment-bound forms of IGF-I with [N29,S39_PQ]IGF-II. **A**, an overlay of the backbone of free human IGF-I (Protein Data Bank code 1GZR; in blue) with [N29,S39_PQ]IGF-II (in purple) and IGF-I from a complex with IR and IGF-1R α -CT peptide (Protein Data Bank code 4XSS; in cyan). The positions of Asn²⁶ in IGF-I and Asn²⁹ in IGF-II side chains are highlighted. **B**, the crystal structure (Protein Data Bank code 4XSS) of IGF-I (in cyan) in a complex with IR L1 domain (in white) and IGF-1R α -CT peptide (in green) overlaid with [N29,S39_PQ]IGF-II in purple.

[N29,S39_PQ]IGF-II) may be involved in direct contacts to IGF-1R, and this hypothesis could be supported by a positive effect of S29N mutation on IGF-1R binding affinity of [N29,S39_PQ]IGF-II. Hence, Ser²⁹ may represent an important site for engineering of the IGF-1R binding specificity in IGF-II analogs.

Concluding Remarks

We have developed a straightforward protocol for the production of recombinant IGF-II with an additional glycine at the N terminus. We prepared six IGF-II analogs with IGF-I-like mutations. All these mutants have markedly reduced affinity for IR-A, especially those analogs with Pro-Gln insertions in the C-domain. Moreover, one of the analogs, [N29,S39_PQ]IGF-II, shows the enhanced binding affinity for IGF-1R in comparison with IGF-II due to the synergistic effect of Pro-Gln insertion and S29N point mutation. Consequently, this analog has almost 10-fold enhanced IGF-1R/IR-A binding selectivity in comparison with IGF-II. Structural characterization of selected analogs revealed that the conformational rearrangement of the C-loop induced by insertion of two residues from IGF-I is manifested in the reduced affinity for IR-A. A combination of the effect of this insertion with an additional IGF-I like substitution, S29N, driving the additional subtle rearrangement of the C-loop forms a structural basis for the increased binding affinity of [N29,S39_PQ]IGF-II for IGF-1R. To our knowledge, the research reported here is a unique example of the determination of 3D structures of IGF-II analogs with modifications that have an impact on receptor binding affinities. Identification of structural determinants in IGFs and insulin that are responsible for specific binding to their cognate receptors is important for designing new, more specific hormone analogs with potential therapeutic applications.

Experimental Procedures

Recombinant Expression of IGF-II Analogs

The human IGF-II sequence was cloned into a modified pRSFDuet-1 expression vector fused with an N-terminal His₆

tag, GB1 protein, and TEV protease cleavage site (Glu-Asn-Leu-Tyr-Phe-Gln ↓ Gly). An additional N-terminal Gly (−1) was incorporated to facilitate TEV cleavage. Mutation S29N, Gly-Ser insertion following Arg³⁴, Pro-Gln insertion following Ser³⁹, and a combination of both insertions were obtained by site-directed mutagenesis (QuikChange kit, Agilent Technologies) performed with appropriate mutagenic primers of the IGF-II sequence subcloned into the pBluescript vector. After sequence verification, the mutant fragments were reintroduced into the full-length IGF-II in the expression vector. Constructs were transformed into *E. coli* BL21(λDE3) and cultivated using LB medium or minimal medium containing [¹⁵N]ammonium sulfate and D-[¹³C]glucose. The bacterial culture was grown at 37 °C to an optical density (550 nm) of ~1, induced with 1 mM isopropyl β -D-1-thiogalactopyranoside, and further cultured for 4–5 h. Cells were harvested by centrifugation for 20 min at 4,000 \times g, and cell pellets were stored at −20 °C prior to further processing.

Isolation of Inclusion Bodies

Cells pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM EDTA, 50 μ M PMSF) using 10 ml of buffer/1 g of biomass and homogenized by three passes through an Avestin EmulsiFlex-C3® apparatus at 4 °C and homogenization pressure of 1,200 megapascals. Inclusion bodies from the cell lysate were obtained by centrifugation at 20,000 \times g at 4 °C for 20 min and further washed as a suspension in a wash buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM EDTA) with 0.1% (v/v) Triton X-100, sonicated in an ice bath, and centrifuged (20,000 \times g, 4 °C, 20 min). The wash procedure was repeated in the absence of 0.1% (v/v) Triton X-100, and wet paste consisting of inclusion bodies was stored at −20 °C.

Purification of IGF-II and Analogs

The inclusion bodies were resuspended in a minimum volume (2 ml/g of wet paste) of 50 mM Tris-HCl, pH 8.0 buffer with 300 mM NaCl and sufficient β -mercaptoethanol to yield a final

Receptor Specificity of IGF-II Analogs

concentration of 0.02% (v/v) after the following dilution step. The suspension was gently diluted into 50 mM Tris-HCl, pH 8.0 buffer with 300 mM NaCl and 8 M urea to a final concentration of ~1 g (wet weight of inclusion bodies)/50 ml and incubated for 2–3 h at room temperature with moderate stirring. The solution of the denatured fusion protein was then loaded onto an equilibrated HisTrap HP (5 ml) column connected to an ÄKTA FPLC® system (GE Healthcare), and after washing with 50 mM Tris-HCl, pH 8.0 buffer with 300 mM NaCl, the retained protein was eluted using a 0–500 mM imidazole gradient in 50 mM Tris-HCl, pH 8.0 buffer with 300 mM NaCl within 10 column volumes. The presence of the fusion protein in collected fractions was verified by SDS-PAGE and anti-His₆ Western blotting, and the pooled fractions were dialyzed at 6 °C against 50 mM Tris-HCl, pH 8.0, 300 mM NaCl. The fusion partner was subsequently cleaved by an overnight TEV digestion in the presence of reduced and oxidized glutathione (1.5 mM GSH and 0.15 mM GSSG) at room temperature. Cleaved IGF-II was separated from the fusion protein by a gravity flow nickel chelating chromatography (HIS-Select Nickel Affinity Gel, Sigma-Aldrich) and further desalted on a Chromabond C₄ column (Macherey-Nagel) using 80% CH₃CN (v/v) with 0.1% TFA (v/v) for elution. The collected protein fraction was lyophilized; resuspended in 7% (v/v) acetic acid, 27% (v/v) CH₃CN, 0.03% TFA (v/v); and purified on a semipreparative RP-HPLC column (Vydac 214TP510-C4, 250 × 10 mm) using a CH₃CN/H₂O gradient supplemented with 0.1% TFA (v/v). The separated fractions were lyophilized, the purity of products was analyzed by analytical RP-HPLC, and the identity of the products was verified by high resolution electrospray ionization mass spectrometry (LTQ Orbitrap XL, Thermo Fisher Scientific, Waltham, MA).

NMR Spectroscopy

All NMR data for free IGF-II and analogs were acquired at 25 °C using 600- and 850-MHz Bruker Avance II spectrometers, both of which were equipped with ¹H/¹³C/¹⁵N cryoprobes. To confirm the correct fold of IGF-II analogs, 1D ¹H spectra (unlabeled samples) and 2D ¹H-¹⁵N HSQC spectra were acquired. The NMR spectra were collected using 350- μ l samples of protein (75–380 μ M) dissolved in 50 mM *d*₄-acetic acid (pH 3.0), 5% D₂O (v/v), 0.01% (w/v) NaN₃. Data for IGF-II and analogs bound to a high affinity Domain 11 variant of IGF-2R (D11) (65, 72) were acquired from 350- μ l samples of 200–400 μ M IGF-II·D11 complex in acetate buffer (20 mM *d*₄-acetic acid, pH 4.2, 5% D₂O (v/v), 0.01% (w/v) NaN₃) at 35 °C.

To determine the structure of either free or bound IGF-IIs, a series of double and triple resonance spectra (79, 80) were recorded on ¹³C/¹⁵N uniformly labeled IGF-II or analogs to determine essentially complete sequence-specific resonance backbone and side chain assignments. Constraints for ¹H-¹H distances were derived from 3D ¹⁵N-¹H NOESY-HSQC and ¹³C-¹H NOESY-HMQC, which were acquired using an NOE mixing time of 100 ms.

The family of converged structures was initially calculated using Cyana 2.1 (81). The combined automated NOE assignment and structure determination protocol was used to automatically assign the NOE cross-peaks identified in NOESY spectra and to produce preliminary structures. In addition,

backbone torsion angle constraints, generated from assigned chemical shifts using the program TALOS+ (82), were included in the calculations. Subsequently, five cycles of simulated annealing combined with redundant dihedral angle constraints were used to produce sets of converged structures with no significant restraint violations (distance and van der Waals violations <0.2 Å and dihedral angle constraint violation <5°), which were further refined in explicit solvent using YASARA software with the YASARA force field (83). The structures with the lowest total energy were selected. Analysis of the family of structures obtained was carried out using the Protein Structure Validation Software suite (Northeast Structural Genomics consortium) and MOLMOL (84). The statistics for the resulting structures are summarized in supplemental Table S1.

Circular Dichroism

CD spectra were measured in a quartz cuvette with an optical path length of 0.5 mm (Starna Cells) using a J-815 spectropolarimeter (Jasco, Japan) at room temperature. The far- and near-UV CD spectra were used to identify changes in protein secondary and tertiary structures. The spectral regions were 200–300 nm. The final spectra were obtained as an average of five accumulations. The spectra were corrected for the baseline by subtracting the spectra of the corresponding polypeptide-free solution. Analogs or IGF-II was dissolved and measured in 5% aqueous acetic acid (0.33 mg/ml; 45 μ M).

Receptor Binding Studies

Commercial human insulin and IGF-II were provided by Sigma-Aldrich, and human IGF-I was provided by Tercica.

Human IM-9 Lymphocytes (Human IR-A Isoform)

Receptor binding studies with the insulin receptor in membranes of human IM-9 lymphocytes (containing only human IR-A isoform) were carried out, and *K_d* values were determined according to the procedure described recently (85). Binding data were analyzed by Excel algorithms especially developed for the IM-9 cell system in the laboratory of Prof. Pierre De Meyts (developed by A. V. Groth and R. M. Shymko, Hagedorn Research Institute, Denmark; a kind gift of P. De Meyts) using a method of non-linear regression and a one-site fitting program and taking into account potential depletion of free ligand. Each binding curve was determined in duplicate, and the final dissociation constant (*K_d*) of an analog was calculated from at least three (*n* ≥ 3) independently determined binding curves. The dissociation constant of human ¹²⁵I-insulin was set to 0.3 nM.

Mouse Embryonic Fibroblasts

Human IR-B Isoform—Receptor binding studies with the insulin receptor in membranes of mouse embryonic fibroblasts derived from IGF-I receptor knock-out mice that solely expressed the human IR-B isoform were performed as described in detail previously (86, 87). Binding data were analyzed, and the dissociation constant (*K_d*) was determined with GraphPad Prism 5 software using a method of non-linear regression and a one-site fitting program and taking into account potential depletion of free ligand. *K_d* values of analogs were determined and calculated by the same procedure as for IR-A.

Human IGF-1R—Receptor binding studies with the IGF-I receptor in membranes of mouse embryonic fibroblasts derived from IGF-1R knock-out mice and transfected with human IGF-1R were performed as described previously (86, 87). Binding data were analyzed, and the dissociation constants were determined and calculated by the same method as for IR-B. The dissociation constant of human ^{125}I -IGF-I was set to 0.2 nM. Mouse embryonic fibroblasts expressing human IR-B or IGF-1R were a kind gift from Prof. Antonino Belfiore (University of Magna Grecia, Catanzaro, Italy) and Prof. Renato Baserga (Thomas Jefferson University, Philadelphia, PA). Here we should note that the use of bovine serum albumin (e.g. Sigma-Aldrich A6003) void of “IGF-binding-like” proteins, which interfere with these binding assays, is essential for the preparation of the binding buffer (88).

Author Contributions—R. H. and K. Křížková contributed equally to the paper. R. H., K. Křížková, K. Kedrová, and I. S. carried out protein expression and purification. R. H., P. S., Z. T., and V. V. carried out NMR experiments and structure refinement. K. Křížková, K. Kedrová, M. C., and L. Ž. tested the analogs. M. F. and C. W. carried out DNA cloning. P. U. measured CD spectra. J. J. and L. Ž. conceived the study, designed experiments, and analyzed data. R. H., K. Křížková, M. P. C., J. J., V. V., and L. Ž. wrote the paper. All authors discussed the results and commented on the manuscript.

Acknowledgment—We thank Prof. Marie Urbanová from the University of Chemistry and Technology in Prague for assistance with measuring CD spectra of analogs.

References

- Pandini, G., Frasca, F., Mineo, R., Sciacca, L., Vigneri, R., and Belfiore, A. (2002) Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved. *J. Biol. Chem.* **277**, 39684–39695
- Křížková, K., Chrudinová, M., Povalová, A., Selicharová, I., Collinsová, M., Vaněk, V., Brzozowski, A. M., Jiráček, J., and Žaková, L. (2016) Insulin-like growth factors hybrids as molecular probes of hormone:receptor binding specificity. *Biochemistry* **55**, 2903–2913
- Lee, J., and Pilch, P. F. (1994) The insulin receptor: structure, function, and signaling. *Am. J. Physiol. Cell Physiol.* **266**, C319–C334
- Boucher, J., Tseng, Y. H., and Kahn, C. R. (2010) Insulin and insulin-like growth factor-1 receptors act as ligand-specific amplitude modulators of a common pathway regulating gene transcription. *J. Biol. Chem.* **285**, 17235–17245
- Siddle, K. (2012) Molecular basis of signaling specificity of insulin and IGF receptors: neglected corners and recent advances. *Front. Endocrinol.* **3**, 34
- Hers, L., Vincent, E. E., and Tavaré, J. M. (2011) Akt signalling in health and disease. *Cell. Signal.* **23**, 1515–1527
- Siddle, K. (2011) Signalling by insulin and IGF receptors: supporting acts and new players. *J. Mol. Endocrinol.* **47**, R1–R10
- Bedinger, D. H., and Adams, S. H. (2015) Metabolic, anabolic, and mitogenic insulin responses: a tissue-specific perspective for insulin receptor activators. *Mol. Cell. Endocrinol.* **415**, 143–156
- Esposito, D. L., Blakesley, V. A., Koval, A. P., Scrimgeour, A. G., and LeRoith, D. (1997) Tyrosine residues in the C-terminal domain of the insulin-like growth factor-I receptor mediate mitogenic and tumorigenic signals. *Endocrinology* **138**, 2979–2988
- O'Connor, R., Kauffmann-Zeh, A., Liu, Y., Lehar, S., Evan, G. I., Baserga, R., and Blättler, W. A. (1997) Identification of domains of the insulin-like growth factor I receptor that are required for protection from apoptosis. *Mol. Cell. Biol.* **17**, 427–435
- Sacco, A., Morcavallo, A., Pandini, G., Vigneri, R., and Belfiore, A. (2009) Differential signaling activation by insulin and insulin-like growth factors I and II upon binding to insulin receptor isoform A. *Endocrinology* **150**, 3594–3602
- LeRoith, D. (1997) Seminars in medicine of the Beth Israel Deaconess Medical Center. Insulin-like growth factors. *N. Engl. J. Med.* **336**, 633–640
- LeRoith, D., and Roberts, C. T. (2003) The insulin-like growth factor system and cancer. *Cancer Lett.* **195**, 127–137
- Dynkevich, Y., Rother, K. I., Whitford, L., Qureshi, S., Galiveeti, S., Szulc, A. L., Danoff, A., Breen, T. L., Kaviani, N., Shanik, M. H., Leroith, D., Vigneri, R., Koch, C. A., and Roth, J. (2013) Tumors, IGF-2, and hypoglycemia: insights from the clinic, the laboratory, and the historical archive. *Endocr. Rev.* **34**, 798–826
- Alvino, C. L., Ong, S. C., McNeil, K. A., Delaine, C., Booker, G. W., Wallace, J. C., and Forbes, B. E. (2011) Understanding the mechanism of insulin and insulin-like growth factor (IGF) receptor activation by IGF-II. *PLoS One* **6**, e27488
- Gallagher, E. J., and LeRoith, D. (2011) Minireview: IGF, insulin, and cancer. *Endocrinology* **152**, 2546–2551
- Alberini, C. M., and Chen, D. Y. (2012) Memory enhancement: consolidation, reconsolidation and insulin-like growth factor 2. *Trends Neurosci.* **35**, 274–283
- Chen, D. Y., Stern, S. A., Garcia-Osta, A., Saunier-Rebori, B., Pollonini, G., Bambah-Mukku, D., Blitzer, R. D., and Alberini, C. M. (2011) A critical role for IGF-II in memory consolidation and enhancement. *Nature* **469**, 491–497
- Pascual-Lucas, M., Viana da Silva, S., Di Scala, M., Garcia-Barroso, C., González-Aseguinolaza, G., Mülle, C., Alberini, C. M., Cuadrado-Tejedor, M., and Garcia-Osta, A. (2014) Insulin-like growth factor 2 reverses memory and synaptic deficits in APP transgenic mice. *EMBO Mol. Med.* **6**, 1246–1262
- Clemmons, D. R. (1998) Role of insulin-like growth factor binding proteins in controlling IGF actions. *Mol. Cell. Endocrinol.* **140**, 19–24
- Firth, S. M., and Baxter, R. C. (2002) Cellular actions of the insulin-like growth factor binding proteins. *Endocr. Rev.* **23**, 824–854
- Kornfeld, S. (1992) Structure and function of the mannose 6-phosphate insulin-like growth factor-II receptors. *Annu. Rev. Biochem.* **61**, 307–330
- Belfiore, A., and Malaguarnera, R. (2011) Insulin receptor and cancer. *Endocr.-Relat. Cancer* **18**, R125–R147
- Schaffer, M. L., Deshayes, K., Nakamura, G., Sidhu, S., and Skelton, N. J. (2003) Complex with a phage display-derived peptide provides insight into the function of insulin-like growth factor I. *Biochemistry* **42**, 9324–9334
- Laajoki, L. G., Francis, G. L., Wallace, J. C., Carver, J. A., and Keniry, M. A. (2000) Solution structure and backbone dynamics of long-[Arg³]insulin-like growth factor-I. *J. Biol. Chem.* **275**, 10009–10015
- De Wolf, E., Gill, R., Geddes, S., Pitts, J., Wollmer, A., and Grötzinger, J. (1996) Solution structure of a mini IGF-1. *Protein Sci.* **5**, 2193–2202
- Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T., and Kobayashi, Y. (1993) 3-Dimensional structure of human insulin-like growth factor-I (IGF-I) determined by ¹H-NMR and distance geometry. *Int. J. Pept. Protein Res.* **41**, 433–440
- Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991) Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study. *Biochemistry* **30**, 5484–5491
- Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenburg, J. P., de Bree, F. M., and Dauter, Z. (2002) Structural origins of the functional divergence of human insulin-like growth factor-I and insulin. *Biochemistry* **41**, 9389–9397
- Siwanowicz, I., Popowicz, G. M., Wisniewska, M., Huber, R., Kuenkele, K. P., Lang, K., Engh, R. A., and Holak, T. A. (2005) Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. *Structure* **13**, 155–167
- Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001) Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions. *Biochemistry* **40**, 11022–11029
- Yun, C. H., Tang, Y. H., Feng, Y. M., An, X. M., Chang, W. R., and Liang, D. C. (2005) 1.42 Å crystal structure of mini-IGF-1(2): an analysis of the

Receptor Specificity of IGF-II Analogs

- disulfide isomerization property and receptor binding property of IGF-1 based on the three-dimensional structure. *Biochem. Biophys. Res. Commun.* **326**, 52–59
33. Sitar, T., Popowicz, G. M., Siwanowicz, I., Huber, R., and Holak, T. A. (2006) Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13028–13033
34. Zeslawski, W., Beisel, H. G., Kamionka, M., Kalus, W., Engh, R. A., Huber, R., Lang, K., and Holak, T. A. (2001) The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5. *EMBO J.* **20**, 3638–3644
35. Terasawa, H., Kohda, D., Hatanaka, H., Nagata, K., Higashihashi, N., Fujiwara, H., Sakano, K., and Inagaki, F. (1994) Solution structure of human insulin-like growth-factor-II; recognition sites for receptors and binding-proteins. *EMBO J.* **13**, 5590–5597
36. Torres, A. M., Forbes, B. E., Aplin, S. E., Wallace, J. C., Francis, G. L., and Norton, R. S. (1995) Solution structure of human insulin-like growth-factor-II. Relationship to receptor and binding-protein interactions. *J. Mol. Biol.* **248**, 385–401
37. Gursky, O., Li, Y., Badger, J., and Caspar, D. L. (1992) Monovalent cation binding to cubic insulin crystals. *Biophys. J.* **61**, 604–611
38. McKern, N. M., Lawrence, M. C., Streltsov, V. A., Lou, M. Z., Adams, T. E., Lovrecz, G. O., Elleman, T. C., Richards, K. M., Bentley, J. D., Pilling, P. A., Hoyne, P. A., Cartledge, K. A., Pham, T. M., Lewis, J. L., Sankovich, S. E., et al. (2006) Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature* **443**, 218–221
39. Lawrence, M. C., McKern, N. M., and Ward, C. W. (2007) Insulin receptor structure and its implications for the IGF-1 receptor. *Curr. Opin. Struct. Biol.* **17**, 699–705
40. Ward, C. W., Menting, J. G., and Lawrence, M. C. (2013) The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation. *BioEssays* **35**, 945–954
41. Yamaguchi, Y., Flier, J. S., Benecke, H., Ransil, B. J., and Moller, D. E. (1993) Ligand-binding properties of the two isoforms of the human insulin receptor. *Endocrinology* **132**, 1132–1138
42. Seino, S., Seino, M., Nishi, S., and Bell, G. I. (1989) Structure of the human insulin receptor gene and characterization of its promoter. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 114–118
43. Mosthaf, L., Grako, K., Dull, T. J., Coussens, L., Ullrich, A., and McClain, D. A. (1990) Functionally distinct insulin-receptors generated by tissue-specific alternative splicing. *EMBO J.* **9**, 2409–2413
44. Schaefer, E. M., Siddle, K., and Ellis, L. (1990) Deletion analysis of the human insulin receptor ectodomain reveals independently folded soluble subdomains and insulin binding by a monomeric α -subunit. *J. Biol. Chem.* **265**, 13248–13253
45. Brandt, J., Andersen, A. S., and Kristensen, C. (2001) Dimeric fragment of the insulin receptor α -subunit binds insulin with full holoreceptor affinity. *J. Biol. Chem.* **276**, 12378–12384
46. De and Meyts, P. (2015) Insulin/receptor binding: the last piece of the puzzle? *BioEssays* **37**, 389–397
47. Kristensen, C., Kjeldsen, T., Wiberg, F. C., Schäffer, L., Hach, M., Have-lund, S., Bass, J., Steiner, D. F., and Andersen, A. S. (1997) Alanine scanning mutagenesis of insulin. *J. Biol. Chem.* **272**, 12978–12983
48. Denley, A., Cosgrove, L. J., Booker, G. W., Wallace, J. C., and Forbes, B. E. (2005) Molecular interactions of the IGF system. *Cytokine Growth Factor Rev.* **16**, 421–439
49. Menting, J. G., Whittaker, J., Margetts, M. B., Whittaker, L. J., Kong, G. K., Smith, B. J., Watson, C. J., Záková, L., Kletvíková, E., Jiráček, J., Chan, S. J., Steiner, D. F., Dodson, G. G., Brzozowski, A. M., Weiss, M. A., et al. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241–245
50. Menting, J. G., Yang, Y., Chan, S. J., Phillips, N. B., Smith, B. J., Whittaker, J., Wickramasinghe, N. P., Whittaker, L. J., Pandeyarajan, V., Wan, Z. L., Yadav, S. P., Carroll, J. M., Strokes, N., Roberts, C. T., Jr., Ismail-Beigi, F., et al. (2014) Protective hinge in insulin opens to enable its receptor engagement. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3395–E3404
51. Jiráček, J., Záková, L., Antolíková, E., Watson, C. J., Turkenburg, J. P., Dodson, G. G., and Brzozowski, A. M. (2010) Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1966–1970
52. Záková, L., Kletvíková, E., Veverka, V., Lepšík, M., Watson, C. J., Turkenburg, J. P., Jiráček, J., and Brzozowski, A. M. (2013) Structural integrity of the B24 site in human insulin is important for hormone functionality. *J. Biol. Chem.* **288**, 10230–10240
53. Záková, L., Kletvíková, E., Lepšík, M., Collinsová, M., Watson, C. J., Turkenburg, J. P., Jiráček, J., and Brzozowski, A. M. (2014) Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2765–2774
54. Menting, J. G., Lawrence, C. F., Kong, G. K., Margetts, M. B., Ward, C. W., and Lawrence, M. C. (2015) Structural congruency of ligand binding to the insulin and insulin/type 1 insulin-like growth factor hybrid receptors. *Structure* **23**, 1271–1282
55. Mayer, J. P., Zhang, F., and DiMarchi, R. D. (2007) Insulin structure and function. *Biopolymers* **88**, 687–713
56. Hashimoto, R., Fujiwara, H., Higashihashi, N., Enjoh-Kimura, T., Terasawa, H., Fujita-Yamaguchi, Y., Inagaki, F., Perdue, J. F., and Sakano, K. (1995) N-terminal deletion mutants of insulin-like growth factor-II (IGF-II) show Thr⁷ and Leu⁸ important for binding to insulin and IGF-I receptors and Leu⁸ critical for All IGF-II functions. *J. Biol. Chem.* **270**, 18013–18018
57. Denley, A., Bonython, E. R., Booker, G. W., Cosgrove, L. J., Forbes, B. E., Ward, C. W., and Wallace, J. C. (2004) Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. *Mol. Endocrinol.* **18**, 2502–2512
58. Ziegler, A. N., Chidambaram, S., Forbes, B. E., Wood, T. L., and Levison, S. W. (2014) Insulin-like growth factor-II (IGF-II) and IGF-II analogs with enhanced insulin receptor- α binding affinity promote neural stem cell expansion. *J. Biol. Chem.* **289**, 4626–4633
59. Henderson, S. T., Brierley, G. V., Surinya, K. H., Priebe, I. K., Catcheside, D. E., Wallace, J. C., Forbes, B. E., and Cosgrove, L. J. (2015) Delineation of the IGF-II C domain elements involved in binding and activation of the IR-A, IR-B and IGF-IR. *Growth Horm. IGF Res.* **25**, 20–27
60. Alvino, C. L., McNeil, K. A., Ong, S. C., Delaine, C., Booker, G. W., Wallace, J. C., Whittaker, J., and Forbes, B. E. (2009) A novel approach to identify two distinct receptor binding surfaces of insulin-like growth factor II. *J. Biol. Chem.* **284**, 7656–7664
61. Delaine, C., Alvino, C. L., McNeil, K. A., Mulhern, T. D., Gauguin, L., De Meyts, P., Jones, E. Y., Brown, J., Wallace, J. C., and Forbes, B. E. (2007) A novel binding site for the human insulin-like growth factor-II (IGF-II)/mannose 6-phosphate receptor on IGF-II. *J. Biol. Chem.* **282**, 18886–18894
62. Sakano, K., Enjoh, T., Numata, F., Fujiwara, H., Marumoto, Y., Higashihashi, N., Sato, Y., Perdue, J. F., and Fujita-Yamaguchi, Y. (1991) The design, expression, and characterization of human insulin-like growth factor-II (IGF-II) mutants specific for either the IGF-II cation-independent mannose 6-phosphate receptor or IGF-I receptor. *J. Biol. Chem.* **266**, 20626–20635
63. Zhou, P., and Wagner, G. (2010) Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies. *J. Biomol. NMR* **46**, 23–31
64. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal Protein-G. *Science* **253**, 657–661
65. Williams, C., Hoppe, H. J., Rezugui, D., Strickland, M., Forbes, B. E., Grutzner, F., Frago, S., Ellis, R. Z., Wattana-Amorn, P., Prince, S. N., Zaccaro, O. J., Nolan, C. M., Mungall, A. J., Jones, E. Y., Crump, M. P., et al. (2012) An exon splice enhancer primes IGF2:IGF2R binding site structure and function evolution. *Science* **338**, 1209–1213
66. Frasca, F., Pandini, G., Scalia, P., Sciacca, L., Mineo, R., Costantino, A., Goldfine, I. D., Belfiore, A., and Vigneri, R. (1999) Insulin receptor isoform A, a newly recognized, high-affinity insulin-like growth factor II receptor in fetal and cancer cells. *Mol. Cell. Biol.* **19**, 3278–3288
67. Slaaby, R. (2015) Specific insulin/IGF1 hybrid receptor activation assay reveals IGF1 as a more potent ligand than insulin. *Sci. Rep.* **5**, 7911

68. Cottam, J. M., Scanlon, D. B., Karas, J. A., Calabrese, A. N., Pukala, T. L., Forbes, B. E., Wallace, J. C., and Abell, A. D. (2013) Chemical synthesis of a fluorescent IGF-II analogue. *Int. J. Pept. Res. Ther.* **19**, 61–69
69. King, G. L., Kahn, C. R., Samuels, B., Danho, W., Bullesbach, E. E., and Gattner, H. G. (1982) Synthesis and characterization of molecular hybrids of insulin and insulin-like growth factor-I. The role of the A-chain extension peptide. *J. Biol. Chem.* **257**, 10869–10873
70. Francis, G. L., Ross, M., Ballard, F. J., Milner, S. J., Senn, C., McNeil, K. A., Wallace, J. C., King, R., and Wells, J. R. (1992) Novel recombinant fusion protein analogs of insulin-like growth-factor (IGF)-I indicate the relative importance of IGF-binding protein and receptor-binding for enhanced biological potency. *J. Mol. Endocrinol.* **8**, 213–223
71. Francis, G. L., Aplin, S. E., Milner, S. J., McNeil, K. A., Ballard, F. J., and Wallace, J. C. (1993) Insulin-like growth-factor (IGF)-II binding to IGF-binding proteins and IGF receptors is modified by deletion of the N-terminal hexapeptide or substitution of arginine for glutamate-6 in IGF-II. *Biochem. J.* **293**, 713–719
72. Williams, J., Rezgui, D., Prince, S. N., Zaccheo, O. J., Foulstone, E. J., Forbes, B. E., Norton, R. S., Crosby, J., Hassan, A. B., and Crump, M. P. (2007) Structural insights into the interaction of insulin-like growth factor 2 with IGF2R domain 11. *Structure* **15**, 1065–1078
73. Sohma, Y., Pentelute, B. L., Whittaker, J., Hua, Q. X., Whittaker, L. J., Weiss, M. A., and Kent, S. B. (2008) Comparative properties of insulin-like growth factor I (IGF-I) and [Gly7D-Ala]IGF-1 prepared by total chemical synthesis. *Angew. Chem. Int. Ed. Engl.* **47**, 1102–1106
74. Gill, R., Verma, C., Wallach, B., Urso, B., Pitts, J., Wollmer, A., De Meyts, P., and Wood, S. (1999) Modelling of the disulphide-swapped isomer of human insulin-like growth factor-1: implications for receptor binding. *Protein Eng.* **12**, 297–303
75. Bayne, M. L., Applebaum, J., Underwood, D., Chicchi, G. G., Green, B. G., Hayes, N. S., and Cascieri, M. A. (1989) The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor. *J. Biol. Chem.* **264**, 11004–11008
76. Bayne, M. L., Applebaum, J., Chicchi, G. G., Miller, R. E., and Cascieri, M. A. (1990) The roles of tyrosine-24, tyrosine-31, and tyrosine-60 in the high-affinity binding of insulin-like growth factor-I to the type-1 insulin-like growth-factor receptor. *J. Biol. Chem.* **265**, 15648–15652
77. Maly, P., and Lüthi, C. (1988) The binding sites of insulin-like growth factor I (IGF I) to type I IGF receptor and to a monoclonal antibody. Mapping by chemical modification of tyrosine residues. *J. Biol. Chem.* **263**, 7068–7072
78. Keyhanfar, M., Booker, G. W., Whittaker, J., Wallace, J. C., and Forbes, B. E. (2007) Precise mapping of an IGF-I-binding site on the IGF-1R. *Biochem. J.* **401**, 269–277
79. Renshaw, P. S., Veverka, V., Kelly, G., Frenkiel, T. A., Williamson, R. A., Gordon, S. V., Hewinson, R. G., and Carr, M. D. (2004) Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the *Mycobacterium tuberculosis* proteins CFP-10 and ESAT-6. *J. Biomol. NMR* **30**, 225–226
80. Veverka, V., Lennie, G., Crabbe, T., Bird, I., Taylor, R. J., and Carr, M. D. (2006) NMR assignment of the mTOR domain responsible for rapamycin binding. *J. Biomol. NMR* **36**, Suppl. 1, 3
81. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227
82. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223
83. Harjes, E., Harjes, S., Wohlgemuth, S., Müller, K. H., Krieger, E., Herrmann, C., and Bayer, P. (2006) GTP-Ras disrupts the intramolecular complex of C1 and RA domains of Nore1. *Structure* **14**, 881–888
84. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55, 29–32
85. Morcavallo, A., Genua, M., Palumbo, A., Kletvíková, E., Jiráček, J., Brzozowski, A. M., Iozzo, R. V., Belfiore, A., and Morrione, A. (2012) Insulin and insulin-like growth factor II differentially regulate endocytic sorting and stability of insulin receptor isoform A. *J. Biol. Chem.* **287**, 11422–11436
86. Kosinová, L., Veverka, V., Novotná, P., Collinsová, M., Urbanová, M., Moody, N. R., Turkenburg, J. P., Jiráček, J., Brzozowski, A. M., and Žáková, L. (2014) Insight into the structural and biological relevance of the T/R transition of the N-terminus of the B-chain in human insulin. *Biochemistry* **53**, 3392–3402
87. Viková, J., Collinsová, M., Kletvíková, E., Buděšínský, M., Kaplan, V., Žáková, L., Veverka, V., Hexnerová, R., Tarazona Aviñó, R. J., Straková, J., Selicharová, I., Vaněk, V., Wright, D. W., Watson, C. J., Turkenburg, J. P., et al. (2016) Rational steering of insulin binding specificity by intra-chain chemical crosslinking. *Sci. Rep.* **6**, 19431
88. Slaaby, R., Andersen, A. S., and Brandt, J. (2008) IGF-I binding to the IGF-I receptor is affected by contaminants in commercial BSA: the contaminants are proteins with IGF-I binding properties. *Growth Horm. IGF Res.* **18**, 267–274

Supplementary Information

for

Probing receptor specificity by sampling the conformational space of the insulin-like growth factor II C-domain

Rozálie Hexnerová^{†§#}, Květoslava Křížková^{†§#}, Milan Fábry^{†§}, Irena Siegllová[‡], Kateřina Kedrová^{‡§}, Michaela Collinsová[‡], Pavlína Ullrichová[†], Pavel Srb[‡], Christopher Williams[£], Matthew P. Crump[£], Zdeněk Tošner[§], Jiří Jiráček[‡], Václav Veverka^{†*} and Lenka Žáková^{†*}

[‡]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic,
v.v.i., Flemingovo nám 2, 166 10 Prague 6, Czech Republic

[§]Faculty of Science, Charles University in Prague, Albertov 6, Prague, 128 43,
Czech Republic

[†]Department of Analytical Chemistry, University of Chemistry and Technology, Prague,
Technická 5, 166 28 Prague 6,
Czech Republic

[§]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, v.v.i., Vídeňská
1083, 142 20 Prague 4, Czech Republic

[£]Department of Organic and Biological Chemistry, School of Chemistry, Cantock's Close,
University of Bristol, Bristol BS8 1TS, United Kingdom

[#]Joint first authors

Table of Contents

Figure S1.....	2
Figure S2.....	3
Figure S3.....	4
Figure S4.....	5
Figure S5.....	6
Figure S6.....	6
Figure S7.....	7
Figure S8.....	8
Table S1.....	9

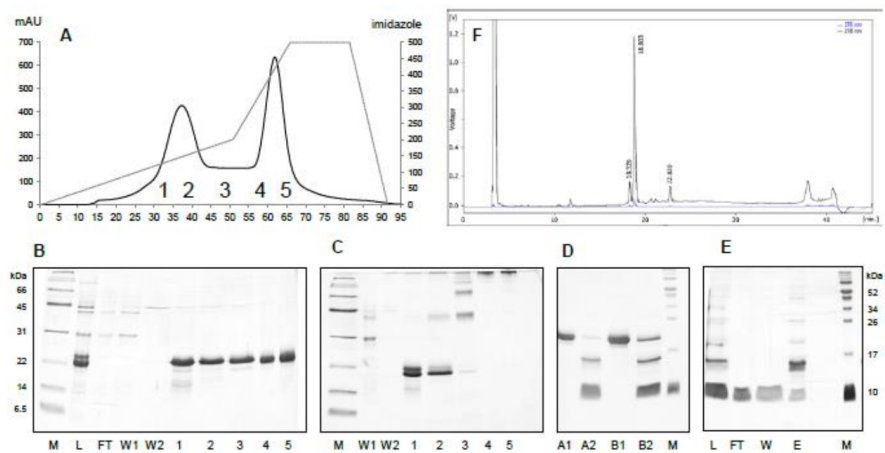


Figure S1. Purification procedure for IGF-II analogs. **A.** The elution profile from purification of denatured IGF-II in fusion with GB1 protein by IMAC. The material eluted in two major fractions (1-2 and 4-5) at two different imidazole concentrations. SDS-PAGE analysis of collected fractions (1-5) under reducing (**B**) and non-reducing (**C**) conditions revealing the presence of two monomeric isoforms (folded and misfolded) eluting at lower concentration of imidazole (150 mM) and multimeric aggregates eluting at higher imidazole concentration (400 mM). M, molecular weight standard; L, sample load; FT, flow through; W1 and W2, wash; 1-5, eluted fractions. Panel **D** shows reducing SDS-PAGE of the fusion partner cleavage by TEV protease. A1, monomeric fractions before TEV addition; A2, monomeric fractions after 24hrs of TEV digestion; B1, multimeric fraction before TEV addition; B2, multimeric fractions after 24hrs of TEV digestion; M, molecular weight standards. Panel **E** shows reducing SDS-PAGE of cleaved sample after nickel chelating chromatography. The cleaved IGF-II is present in FT and W fraction. L, sample load, FT, flow through; W, wash; E, elution; M, molecular weight standard. Panel **F** shows the final RP-HPLC purification of IGF-II separating forms with differently linked disulfide bonds.

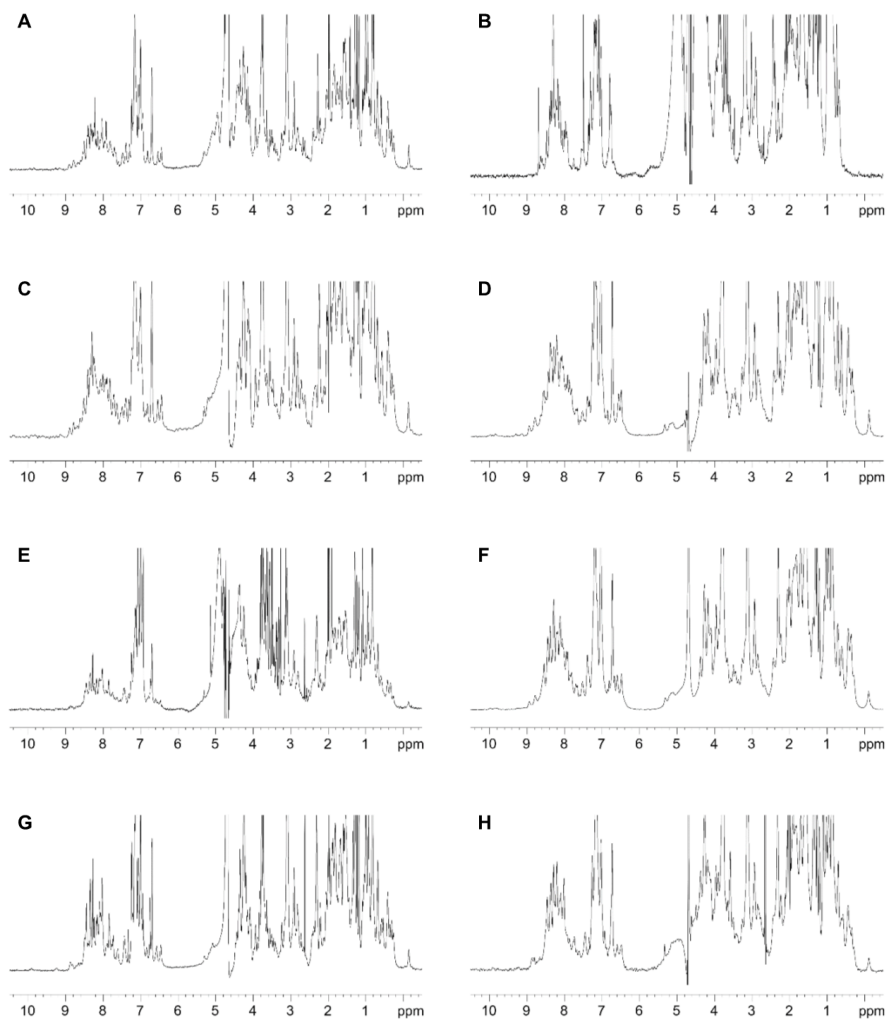


Figure S2. ^1H NMR spectra of IGF-II analogues. (A) IGF-II, (B) misfolded IGF-II, (C) [N29]-IGF-II, (D) [R34_GS]-IGF-II, (E) [S39_PQ]-IGF-II, (F) [R34_GS,S39_PQ]-IGF-II, (G) [N29, S39_PQ]-IGF-II, (H) [N29, R34_GS, S39_PQ]-IGF-II. The difference between correctly folded (A) and misfolded (B) IGF-II spectra was used for verification of correct protein folding of the IGF-II analogs (C-H). In particular, the presence of dispersed aromatic proton signals at 6.5 ppm and upfield shifted methyl signals between 0.5 and -0.2 ppm could be utilized to fingerprint correctly folded IGF-II.

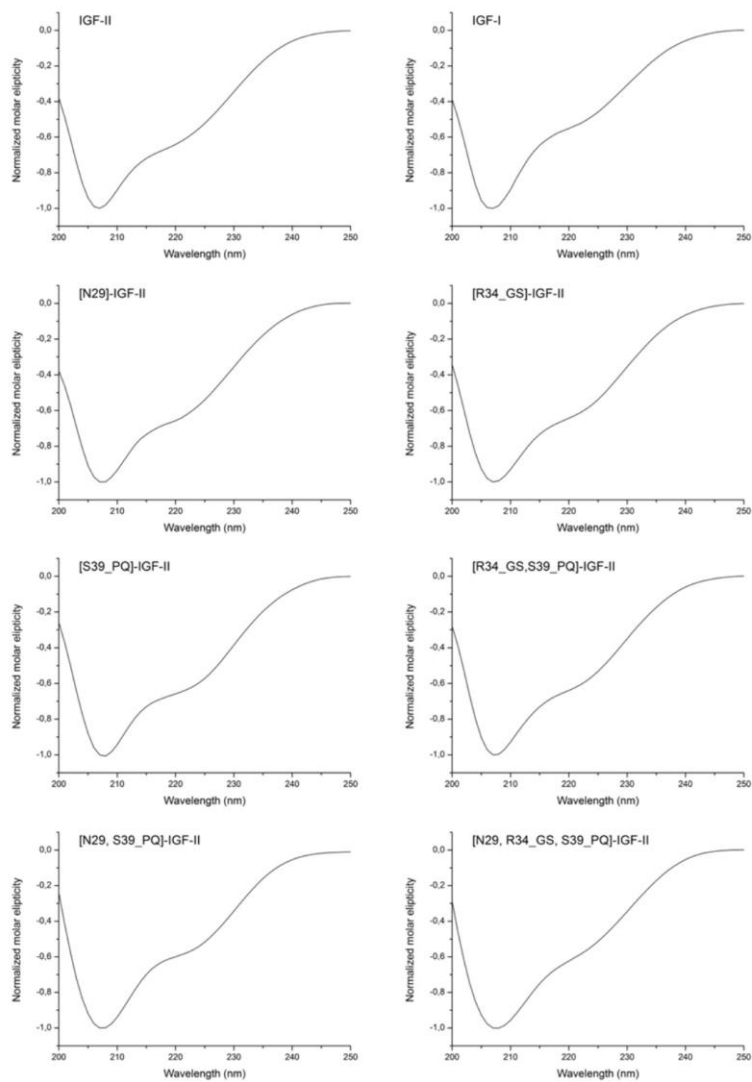


Figure S3. Far UV circular dichroism spectra of IGF-I and studied IGF-II analogs normalized to 207 nm. The curve profiles suggest highly similar presence of the α -helical secondary structure elements in the studied IGF-II analogs.

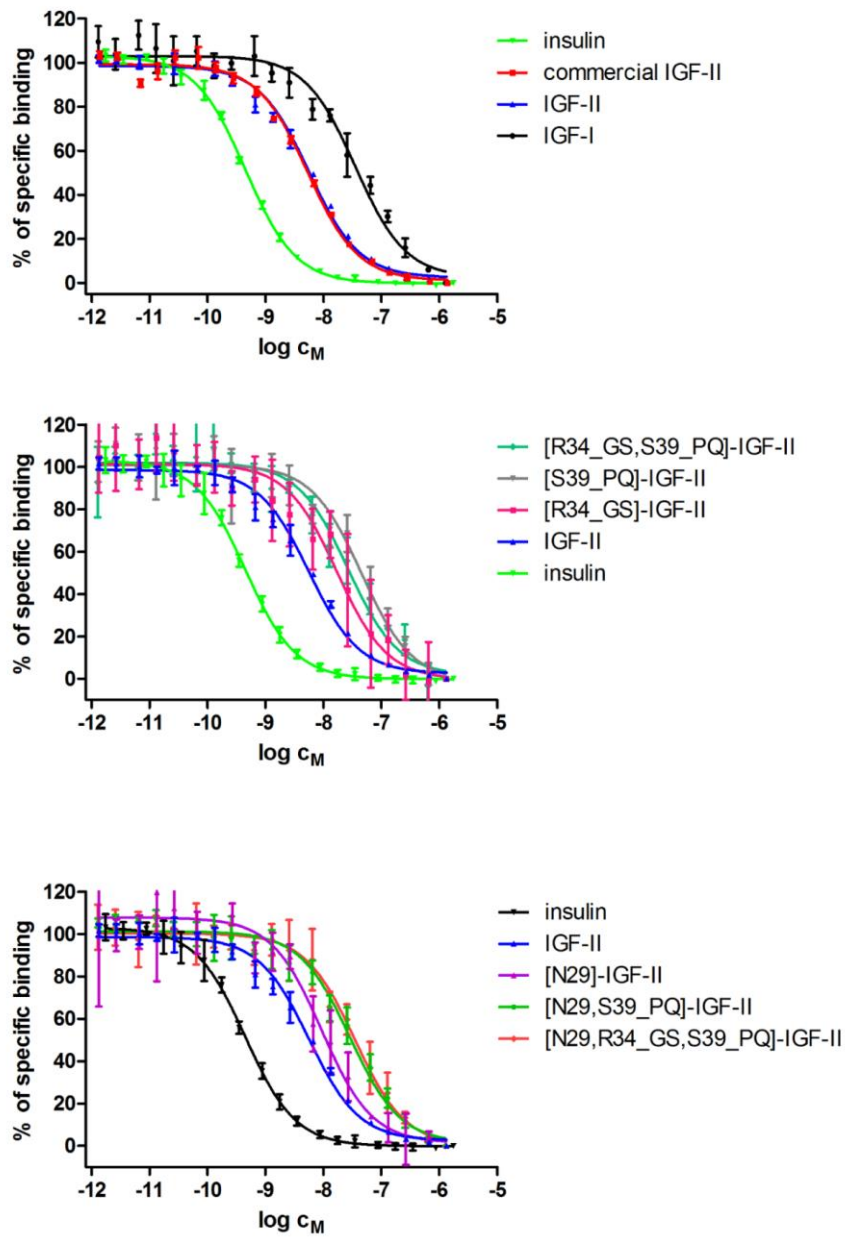


Figure S4. Inhibition of binding of human $[^{125}\text{I}]$ -insulin to IR-A in membranes of IM-9 cells by human insulin, IGF-I, IGF-II and IGF-II analogs.

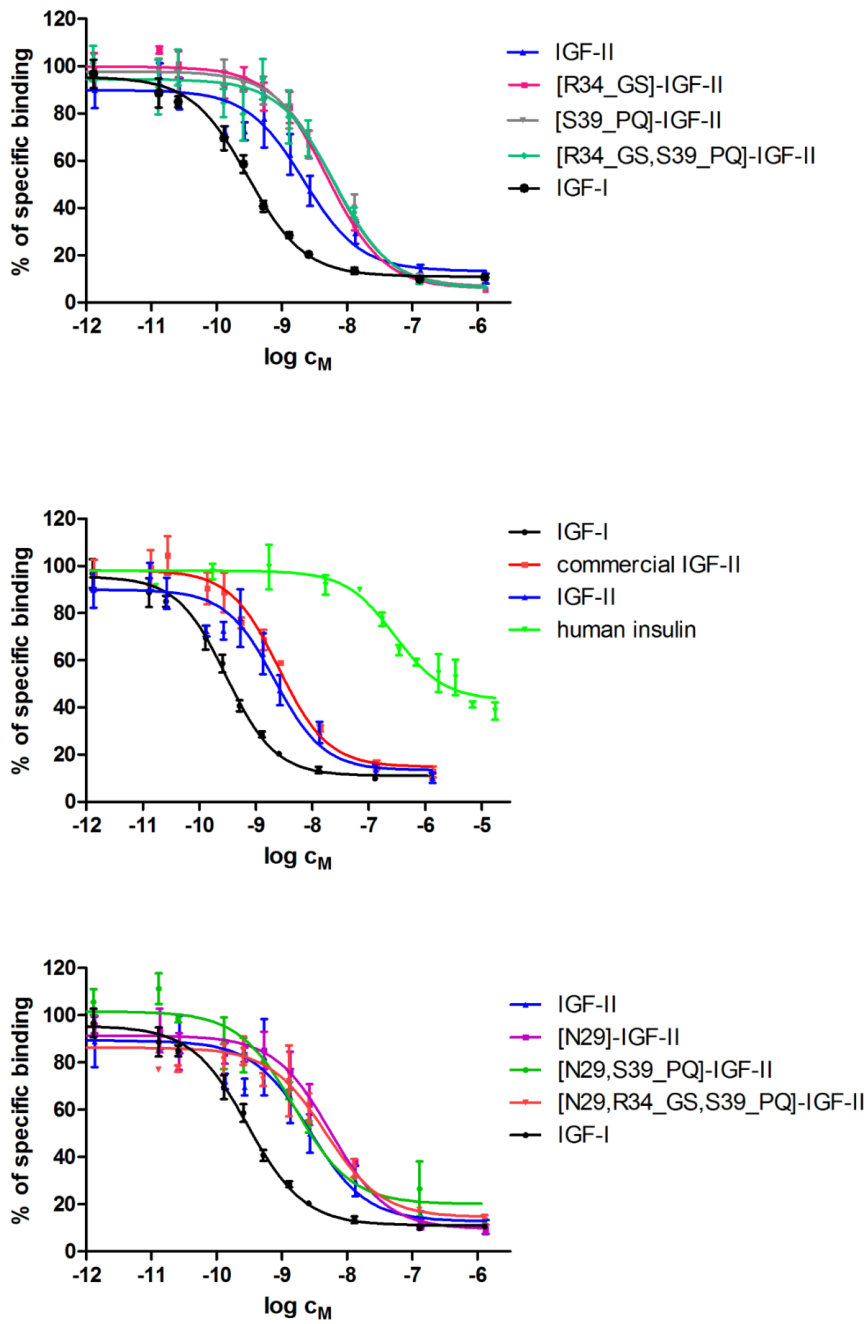


Figure S5. Inhibition of binding of human $[^{125}\text{I}]\text{-IGF-I}$ to IGF-1R in membranes of mouse fibroblasts by human insulin, IGF-I, IGF-II and IGF-II analogs.

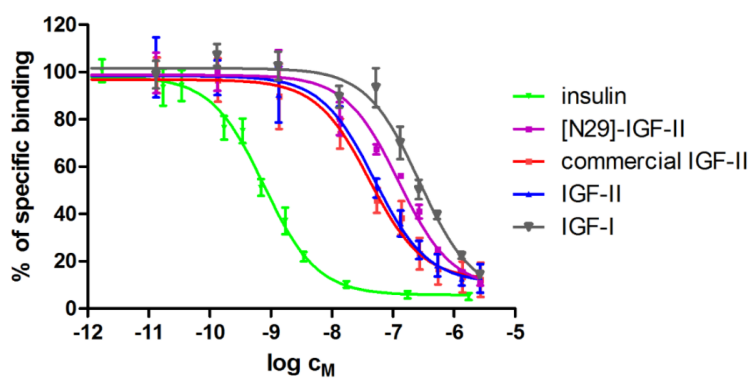


Figure S6. Inhibition of binding of human [¹²⁵I]-insulin to IR-B in membranes of mouse fibroblasts by human insulin, IGF-I, IGF-II and [N29]-IGF-II analog.

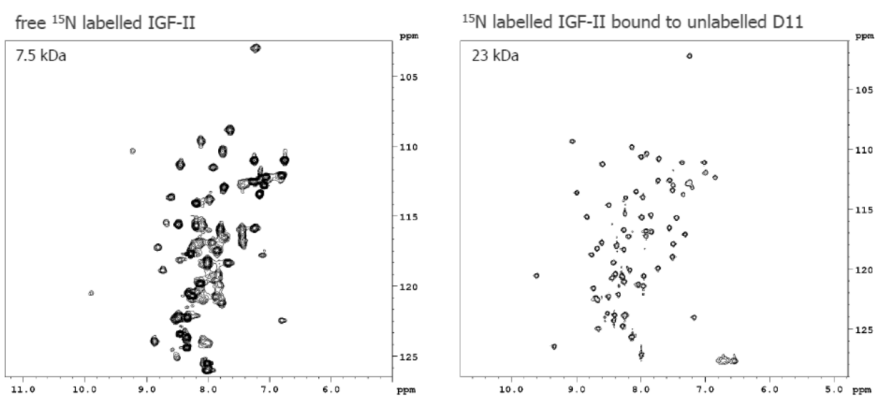


Figure S7. Significant narrowing of IGF-II signals in ¹H/¹⁵N HSQC spectrum upon binding to IGF-2R Domain 11. A spectrum of free ¹⁵N labelled IGF-II is shown on the left panel. Obtained signals do not correspond to the protein mass of 7.5 kDa. The right panel illustrates the signal narrowing observed for IGF-II bound to Domain 11.

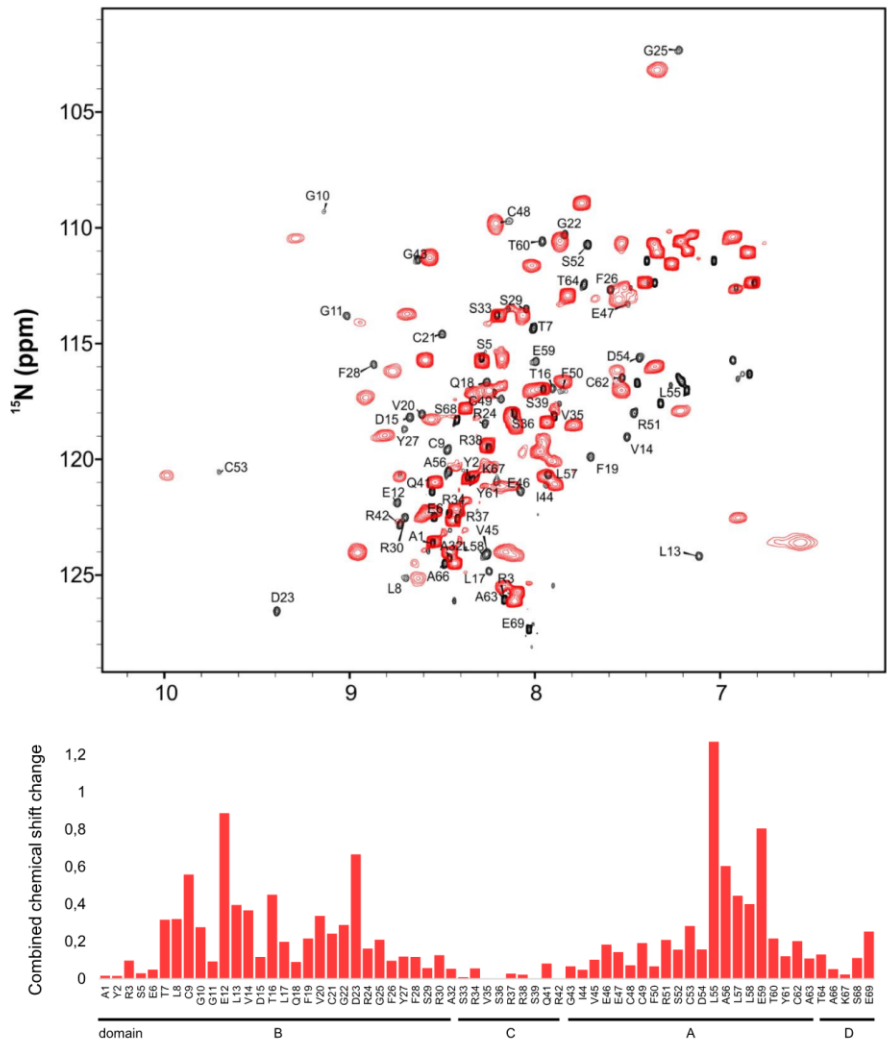


Figure S8. The C-domain of IGF-II is not affected by D11 binding.
 (A) An overlay of $^1\text{H}/^{15}\text{N}$ HSQC spectra obtained for the free (red) and D11-bound [S39_PQ]-IGF-II (black). (B) Values of combined chemical shift changes calculated from the changes of backbone amide signal positions. The major differences upon binding to D11 are distributed across the D11 binding interface, while the signals of the C-domain backbone amides bearing the modifications remain relatively unaffected by the D11 binding.

Table S1. NMR restraints and structural statistics

	<i>IGF-II</i>		<i>[S39_PQ]-IGF-II</i>		<i>[N29, S39_PQ]-IGF-II</i>	
<i>Non-redundant distance and angle constrains</i>						
Total number of NOE constraints	1039		1116		1395	
Short-range NOEs						
Intra-residue (i = j)	301		315		341	
Sequential (i - j = 1)	321		356		406	
Medium-range NOEs (1 < i - j < 5)	160		185		281	
Long-range NOEs (i - j ≥ 5)	254		257		364	
Torsion angles	46		46		46	
Hydrogen bond restrains	-		-		-	
Total number of restricting constraints	1085		1162		1441	
Total restricting constraints per restrained residue	16.2		16.8		20.9	
<i>Residual constraint violations</i>						
Distance violations per structure						
0.1 – 0.2 Å	5.05		5.85		9	
0.2 – 0.5 Å	2.15		2.3		2.6	
> 0.5 Å	0		0		0	
r.m.s. of distance violation per constraint	0.02 Å		0.02 Å		0.02 Å	
Maximum distance violation	0.45 Å		0.48 Å		0.48 Å	
Dihedral angle violations per structure						
1 – 10 °	1.3		1.2		1.7	
> 10 °	0		0		0	
r.m.s. of dihedral violations per constraint	0.68 °		0.71 °		0.75 °	
Maximum dihedral angle violation	5.00 °		5.00 °		5.00 °	
<i>Ramachandran plot summary from Procheck</i>						
Most favoured regions	94.8%		92.2%		85.9%	
Additionally allowed regions	5.2%		7.8%		13.8%	
Generously allowed regions	0.0%		0.0%		0.1%	
Disallowed regions	0.0%		0.0%		0.1%	
<i>r.m.s.d. to the mean structure</i>						
	<i>ordered¹</i>	<i>all</i>	<i>ordered¹</i>	<i>all</i>	<i>ordered¹</i>	<i>all</i>
All backbone atoms	0.4 Å	2.9 Å	1.1 Å	2.2 Å	1.0 Å	1.9 Å
All heavy atoms	1.0 Å	3.6 Å	1.7 Å	2.9 Å	1.4 Å	2.5 Å

¹ Residues with sum of phi and psi order parameters > 1.8

7.4 Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation

Background:

IGF-1 and IGF-2 are protein hormones involved in normal growth and development, but also in abnormal growth like cancer. They exert their activities by binding to different, but highly homologous (ca. 75 %) receptors: IGF-1R, IR-A and IR-B. A high degree of homology of these receptors results in a significant cross-binding of both IGFs, and also insulin, to IR-A and IGF-1R. IGFs mainly bind to the IGF-1R receptor. However, their binding to the isoform A of insulin receptor (IR-A) probably have an important role in the organism too, especially in the case of IGF-2. Both IGF-1 and IGF-2 were shown to be involved in tumorigenesis by binding to IGF-1R and IGF-2 by binding to IR-A as well. The clinical trials with IGF-1R tyrosine kinase inhibitors or specific anti IGF-1R or anti IGF-1 antibodies have not fulfilled expectations (see above). Hence, development of IGF-1R or IR-A antagonists would be of great clinical interest with important clinical applications. Whittaker et al showed that mutations of A4 and A8 sites in human insulin caused a disproportionate effect on hormone IR binding and activation, i.e. partial antagonism [262]. In this study, we were interested if the mutation of respective residues in IGF-1 and IGF-2 can elicit similar effects as in the case of insulin.

Summary:

Here, we prepared recombinant IGF-1 analogues specifically mutated at sites 45, 46 and 49 and IGF-2 analogues at sites 45 and 48, which correspond, or are close, to insulin A4 and A8 sites. Afterwards, we tested the analogues and for their binding and activation of IGF-1R and IR-A binding and their receptor autophosphorylation potencies.

All new analogues retained the main IGF-1R-related properties of native IGF-1 or IGF-2, but the molecules with His49 in IGF-1, and His48 in IGF-2, showed significantly higher affinities for the insulin receptor (for both IR-A and IR-B isoforms) and they are the strongest IGFs-like binders of these receptors ever reported. All hormones activated receptors without major discrepancies with their binding affinities. This study revealed that IR-A and IGF-1R contain specific sites, likely parts of their so-called sites 2', which can interact differently with specifically modified IGF analogues. These findings may facilitate

novel and rational engineering of new hormone analogues for IR-A and IGF-1R studies and for potential medical applications.

My contribution:

I customized the recombinant production protocol of IGF-2 for the production of IGF-1 analogues and I prepared all the IGF-1 analogues in the paper and participated in the measurement of their bindings and signalling properties. I also contributed to the preparation of the manuscript.



Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation

Kateřina Macháčková,[†] Martina Chrudinová,[†] Jelena Radosavljević,^{†,||} Pavlo Potalitsyn,[†] Květoslava Křížková,[†] Milan Fábry,[‡] Irena Selicharová,[†] Michaela Collinsová,[†] Andrzej M. Brzozowski,^{§,®} Lenka Žáková,^{*,†} and Jiří Jiráček^{*,†,®}

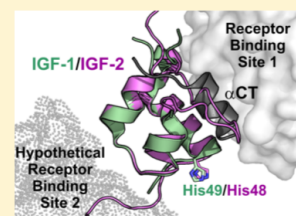
[†]Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo nám 2, 166 10 Prague 6, Czech Republic

[‡]Institute of Molecular Genetics, The Czech Academy of Sciences, Flemingovo n. 2, 166 37 Prague 6, Czech Republic

[§]York Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York YO10 SDD, United Kingdom

Supporting Information

ABSTRACT: Insulin-like growth factors 1 and 2 (IGF-1 and -2, respectively) are protein hormones involved not only in normal growth and development but also in life span regulation and cancer. They exert their functions mainly through the IGF-1R or by binding to isoform A of the insulin receptor (IR-A). The development of IGF-1 and IGF-2 antagonists is of great clinical interest. Mutations of A4 and A8 sites of human insulin lead to disproportionate effects on hormone IR binding and activation. Here, we systematically modified IGF-1 sites 45, 46, and 49 and IGF-2 sites 45 and 48, which correspond, or are close, to insulin sites A4 and A8. The IGF-1R and IR-A binding and autophosphorylation potencies of these analogues were characterized. They retained the main IGF-1R-related properties, but the hormones with His49 in IGF-1 and His48 in IGF-2 showed significantly higher affinities for IR-A and for IR-B, being the strongest IGF-1- and IGF-2-like binders of these receptors ever reported. All analogues activated IR-A and IGF-1R without major discrepancies in their binding affinities. This study revealed that IR-A and IGF-1R contain specific sites, likely parts of their so-called sites 2', which can interact differently with specifically modified IGF analogues. Moreover, a clear importance of IGF-2 site 44 for effective hormone folding was also observed. These findings may facilitate novel and rational engineering of new hormone analogues for IR-A and IGF-1R studies and for potential medical applications.



Two insulin-like growth factors, IGF-1 and IGF-2, together with insulin, are members of a family of small protein hormones that share common evolutionary origins,^{1–4} having similar primary (Figure 1) and three-dimensional structures.⁵ They regulate a wide spectrum of key physiological events, with insulin being responsible mainly for broad, metabolic control,⁶ while IGF-1 and IGF-2 are growth factors involved primarily in the development and growth of mammals.⁵ The role of IGF-1 is relatively well studied,^{7–9} but physiological functions of IGF-2 are much less understood,¹⁰ despite emerging evidence of its impact on the central nervous system.^{11,12}

IGF-1, IGF-2, and insulin exert their activities by binding to different but highly homologous (~75%), ~450 kDa ($\alpha\beta$)₂ dimeric tyrosine-kinase receptors: IGF-1 receptor (IGF-1R) and insulin receptor (IR).^{13–15} IR exists in two isoforms, IR-A, and IR-B, with distinct biochemical properties^{16,17} and a specific tissue distribution. IR-B is a predominant IR form in liver, while muscle and adipose tissues contain both isoforms at different ratios. IR-A is predominant in brain, fetus, and lymphatic tissues and is considered mainly as a “mitogenic” form of the IR, in contrast to IR-B that is considered as the

main “metabolic” receptor for insulin.^{16–18} A high degree of homology of these receptors results in a significant cross-binding of insulin and both IGFs to IR-A and IGF-1R,¹⁹ and hence some overlapping biological responses to binding of these ligands.^{20,21}

The binding of insulin and IGFs to these receptors triggers two major signaling pathways that are initiated by the autophosphorylation of tyrosines within their intracellular tyrosine kinase domains.²² The phosphoinositide 3-kinase (PI3K)/Akt pathway leads to the metabolic, glycemic responses and effects of the hormone:receptor complex, but it is also important for growth and protein synthesis.²³ The Ras/ERK main pathway involves activation of the Ras/Raf/MAPK/ERK1/2 cascade, which mediates proliferative effects through gene transcription regulation.²¹ Whereas insulin signals mainly via both IR isoforms,²⁴ IGF-1 and IGF-2 promote mitogenic

Received: December 15, 2017

Revised: March 29, 2018

Published: April 2, 2018

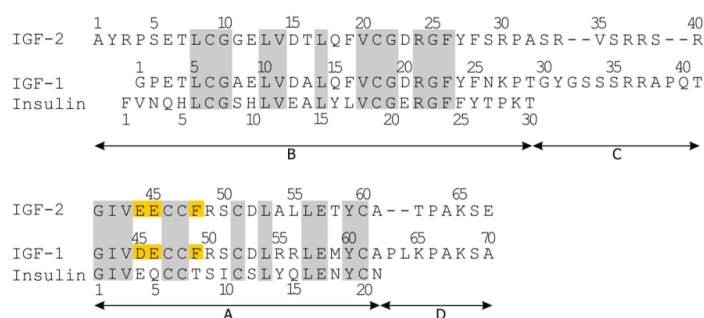


Figure 1. Comparison of the primary sequences of human IGF-1, IGF-2, and insulin. The organization of IGF-1 and IGF-2 into B, C, A, and D domains is shown below the sequences. Insulin A and B chains correspond to IGF A and B domains, respectively. The homologous regions are highlighted in gray, and the residues mutated in this study are highlighted in yellow.

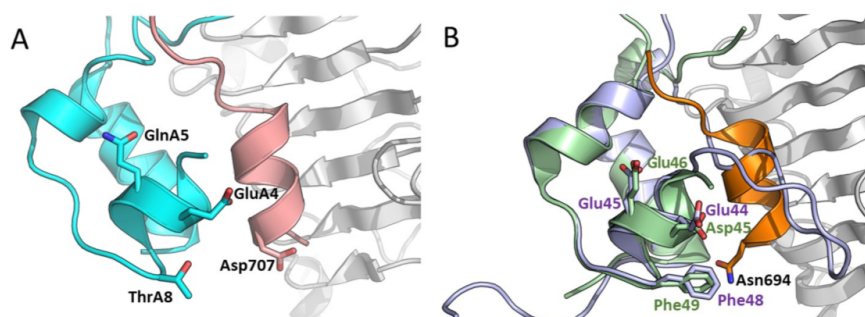


Figure 2. Receptor-bound structures of human insulin and human IGF-1 overlaid with human IGF-2. (A) Complex (4OGA)³⁸ of human insulin sitting on the IR-L1 domain and IR α -CT peptide. Insulin is colored cyan, the L1 domain light gray, and α -CT pink. Side chains of insulin residues GluA4, GlnA5, ThrA8, and α -CT Asp707 are shown and labeled. (B) Complex (4XSS)³⁹ of human IGF-1 (violet) bound to the IR-L1 domain (light gray) and the IGF-1R α -CT peptide (orange). Side chains of IGF-1 residues Asp45, Glu46, and Phe49 mutated in this study and α -CT Asn694 are shown and labeled. The complex is overlaid with the NMR structure of human IGF-2 (5L3L),⁴⁰ which is colored light green. Mutated IGF-2 residues Glu44, Glu45, and Phe48 and α -CT Asp694 are also shown and labeled.

signaling through IGF-1R, and, importantly, similar mitogenic stimulation may result from binding of IGF-2 to IR-A.²⁵ The complexity of insulin/IGF signaling is amplified further by the heterodimerization of IGF-1R and IR-A, and the presence of hybrid receptors that can be effectively activated by IGF-1, but not by insulin.^{26,27} Moreover, the bioavailability of free IGF-1 and IGF-2 for receptor signaling is modulated by a family of high-affinity IGF binding proteins 1–6 (IGFBP 1–6, respectively),^{28,29} and the circulation level of IGF-2 is also affected by a structurally distinct, and presumably nonsignaling, insulin-like growth factor type 2 receptor (IGF-2R), also known as the mannose 6-phosphate receptor.³⁰ The equilibrium of all individual components and the appropriate function of the entire insulin/IGF system are essential for a proper functioning of the organism.³¹

In recent years, the role of the IGF/insulin system in cancer development and growth has been widely studied.^{10,31,32} Substantial efforts have been focused on the development of new anti-IGF-1R-directed therapies, mostly tyrosine-kinase (TK) inhibitors and antireceptor antibodies.³³ However, the results of clinical trials were not satisfactory,³⁴ because of either the toxicity of the TK-targeting drugs or an increasing overlap and takeover of IGF-1R signaling pathways by the IR. The lack of progress in addressing one of the key hallmarks of cancer underlines the need for new anticancer therapies that would

exploit alternative, and specific, targets of the insulin/IGF axis. Here, a high-affinity/no-efficacy IGF-based IGF-1 analogue, i.e., selective antagonist of the IGF-1R, should represent a promising new strategy for combating IGF-1R-related malignancies.

To date, no IGF-like peptide antagonists of the IGF-1R have been identified. However, peptides with good IR/IGF-1R binding and antagonistic properties toward these receptors were discovered by a phage-display technique.^{35,36} Whittaker et al.³⁷ showed that a combination of GluA4His and ThrA8His mutations of human insulin results in insulin analogues with native IR binding affinity but poor efficacy, an impaired ability to stimulate autophosphorylation of IR, and downstream Akt activation. They also proposed that surfaces involving insulin GluA4 and IR Asp707 could be behind the mechanism of receptor activation.

Interestingly, insulin acidic GluA4 is preserved by its equivalent Asp45 and Glu44 in IGF-1 and IGF-2, respectively (Figure 1). However, insulin-neutral Gln A5 is replaced by Glu46, and Glu45 in corresponding sites of IGF-1 and IGF-2. In addition, Asp707 of IR α -CT is replaced by a neutral Asn694 in the IGF-1R α -CT segment (Figure 2).

These correlations and trends between positions A4 and A5 in insulin and their IGFs equivalent 45 and 46 (in IGF-1) and 44 and 55 (in IGF-2) sites prompted us to study (i) the impact

of the mutations of these residues on receptor affinities and potencies of these hormones, and (ii) whether such mutations can generate significant IGF-1R-specific antagonists, with potential anticancer clinical applications. A series of 14 IGF-1 and IGF-2 analogues mutated at these sites have been designed and made, and their binding to IGF-1R and IR-A and abilities to activate the receptors were characterized. Some of these mutations were also combined with the Phe49His mutation in IGF-1 and the Phe48His mutation in IGF-2, as it has been shown that insulin-corresponding mutation ThrA8His significantly increased the IR-A binding potency of this analogue.^{37,41–46}

Despite the results for the analogues studied here not showing any antagonism, they revealed interesting properties of new IGF-1 and IGF-2 mutants that can interact differently with receptors for insulin and IGF-1. This could indicate new directions for a rational engineering of these hormones.

■ MATERIALS AND METHODS

Cloning and Production of IGF-1 and IGF-2 Analogues. As in our previously published research,⁴⁰ both human IGF-1 (UniprotKB entry P05019 amino acids 49–118) and human IGF-2 (UniprotKB entry P01344 amino acids 25–91) have been cloned into a modified pRSFDuet-1 expression vector (kindly provided by E. Bouřa from the Institute of Organic Chemistry and Biochemistry in Prague) as a fusion with an N-terminally His6-tagged GB1 protein and TEV protease cleavage site. An additional N-terminal glycine residue (Gly–1) was incorporated into IGF-1 to enable cleavage by TEV protease (Glu-Asn-Leu-Tyr-Phe-Gln↓Gly). In contrast to our previous study,⁴⁰ the TEV protease cleavage site for the IGF-2 expression construct was modified to Glu-Asn-Leu-Tyr-Phe-Gln↓Ala, cleavage (↓) of which resulted in a native hormone with an N-terminal alanine. [D45H], [D45N], [D45A], [E46H], [E46Q], and [E46A] mutations in IGF-1 analogues were obtained by the standard site-directed mutagenesis protocol (SDM)⁴⁷ using appropriate mutagenic primers (listed in Table S1). After sequence verification, the mutant fragments were reintroduced into the expression vector. Additional [D45N+E46Q], [F49H], [E46H+F49H], [E46Q+F49H], and [D45N+E46Q+F49H] mutations in IGF-1 analogues and in all cloned IGF-2 analogues were introduced either by the overlap-extension polymerase chain reaction (OE strategy in Table S1),⁴⁸ using specific primers as flanking master primers and subsequent recloning into expression vector, or by the standard site-directed mutagenesis as mentioned above (SDM strategy in Table S1), using appropriate mutagenic primers.

We succeeded in expressing and purifying all planned IGF-1 analogues. However, only three IGF-2 analogues (with mutations [E45Q], [F48H], and [E45Q+F48H]) were successfully produced. A list of successfully expressed constructs along with primers used for mutagenesis is provided as Table S1.

All successfully constructed analogues were produced, purified, and characterized, using the procedures described by Hexnerova et al.⁴⁰ The purity of all tested analogues was >95% (and controlled by reverse phase high-performance liquid chromatography analyses and high-resolution mass spectrometry spectra).

Biological Characterization of IGF-1 and IGF-2 Analogues. Binding affinities for the receptors were determined with receptors in the intact cells. Specifically,

binding affinities for IGF-1R were determined with mouse fibroblasts transfected with human IGF-1R and with deleted mouse IGF-1R according to Hexnerova et al.⁴⁰ Binding affinities for IR-A were determined with human IR-A in human IM-9 lymphocytes according to Vikova et al.⁴⁹ Binding affinities for IR-B were determined with mouse fibroblasts transfected with human IR-B and with deleted mouse IGF-1R according to Zakova et al.⁵⁰ Representative binding curves of analogues with the receptors are shown in Figure S1 (IGF-1R), Figure S2 (IR-A), and Figure S3 (IR-B). The binding curve of each analogue was determined in duplicate, and the final dissociation constant (K_d) was calculated from at least three ($n \geq 3$) binding curves (K_d values), determined independently, and compared to binding curves for IGF-1 or for IGF-2, depending on the type of analogue.

The abilities of analogues to induce autophosphorylation of IGF-1R in membranes of mouse fibroblasts transfected with human IGF-1R and with deleted mouse IGF-1R were determined, as described by Machackova et al.⁵¹ The abilities of analogues to induce autophosphorylation of IR-A in mouse fibroblasts transfected with human IR-A and with deleted mouse IGF-1R were determined, as described by Krizkova et al.⁵² Briefly, the cells were stimulated in 24-well plates (Schoeller) (4×10^4 cells per well) after being starved for 4 h in serum-free medium. The cells were stimulated with 10 nM ligand (insulin, IGF-1, IGF-2, or analogues) for 10 min. Stimulation was stopped by snap-freezing. Proteins were routinely analyzed, using immunoblotting and horseradish peroxidase-labeled secondary antibodies (Sigma-Aldrich). The membranes were probed with anti-phospho-IGF-1R β (Tyr1135/1136)/IR β (Tyr1150/1151) (Cell Signaling Technology). The blots were developed using the SuperSignal West Femto maximum sensitivity substrate (Pierce) and analyzed using the ChemiDoc MP Imaging System (Bio-Rad). The autophosphorylation signal density generated by each ligand on a Western blot was expressed as the contribution of phosphorylation relative to the IGF-1 (IGF-1R fibroblasts) respective human insulin (IR-A fibroblasts) signal in the same experiment. Mean standard deviation (SD) values were calculated from four independent experiments ($n = 4$) and compared to those of native IGF-1 or native IGF-2, depending on the type of analogue. A representative example of an immunoblot used for the evaluation of the abilities of analogues to induce autophosphorylation of receptors is shown in Figure S4.

The dose–response curves for human IGF-1 and [His45]-IGF-1, [Asn45]-IGF-1, [Ala45]-IGF-1, [His46]-IGF-1, and [Gln46]-IGF-1 analogues were also measured to determine their EC_{50} values and their abilities to stimulate the autophosphorylation of IGF-1R; here, the same methodology as for the measurements at a single dose (above) was followed. Log(agonist) versus response (variable slope) curve fitting of data was performed with GraphPad Prism 5. The representative curves are shown in the Figure S5. The EC_{50} values (calculated from at least three independent curves) are shown in the Table S2.

The significance of the changes in binding affinities and in the abilities of analogues to stimulate autophosphorylation was calculated using the two-tailed t test.

■ RESULTS

Production of IGF-1 and IGF-2 Analogues. The production of IGF-1 and IGF-2 was achieved by their

recombinant expression in *Escherichia coli* as a fusion with an N-terminal and cleavable His6-tagged GB1 protein (immunoglobulin binding domain B1 of streptococcal protein-G), followed by the cleavage of the fusion protein with TEV protease. Recently, we used this strategy for the synthesis of IGF-2 analogues modified in the hormone's C domain and possessing an extra glycine residue at their N-terminus (position -1).⁴⁰ Here, we modified the TEV cleavage site (see [Materials and Methods](#)) and succeeded in producing native IGF-2 without the additional Gly-1.

However, this strategy was not successful for IGF-1 because of the proline residue at position 2 in IGF-1 ([Figure 1](#)), which hampered TEV protease-mediated cleavage. Therefore, all analogues of the hormone produced in this work have an extra glycine residue (Gly-1) at the protein N-terminus that enabled TEV protease cleavage of the precursor. The presence of Gly-1 did not have any significant effect on the binding properties of the IGF-1 derivative for either tested receptor [IGF-1R or IR-A ([Table 1](#) or [2](#), respectively)], and both native IGF-1 and Gly-1-IGF-1 can be considered as equipotent.

Design of the First Series of Analogues. The first series of analogues was designed with the substitution of IGF-1 Asp45 with "insulin-inspired" His, Asn, and Ala, and a similar strategy was applied for the replacement of IGF-1 Glu46 with His, Gln, and Ala. The Asn45 and Gln46 mutations were also combined. All planned IGF-1 analogues ([Tables 1](#) and [2](#)) were successfully produced in quantities sufficient for their biological and physicochemical characterization.

In parallel, similar mutations were designed for IGF-2 at positions Glu44 and Glu45 that correspond to IGF-1 positions 45 and 46 ([Figure 1](#)). However, only one IGF-2 analogue, [Gln45]-IGF-2, was made with a significant yield that allowed its characterization. All other IGF-2 analogues formed insoluble precipitates after the TEV cleavage/folding steps.

IGF-1R Binding and Activation Properties of IGF-1 Analogues Modified at Positions 45 and 46, and the [Gln45]-IGF-2 Analogue. The analogues were tested for their binding to the IGF-1 receptor (IGF-1R) ([Table 1](#)), and their binding data were compared with the abilities of the analogues to induce autophosphorylation of the IGF-1R at a concentration of 10 nM ([Figure 3A](#)). For human IGF-1 and [His45]-IGF-1, [Asn45]-IGF-1, [Ala45]-IGF-1, [His46]-IGF-1 and [Gln46]-IGF-1 analogues, we also determined EC₅₀ values of their abilities to stimulate autophosphorylation of IGF-1R ([Figure S5](#)).

In general, the mutations did not dramatically alter the binding characteristics of the analogues in comparison with those of native IGFs. However, it can be noted that analogues with mutations at site 45, [His45]-IGF-1, [Asn45]-IGF-1, and [Ala45]-IGF-1, have significantly reduced (29–60%) binding affinities for IGF-1R, with the lowest values being that of [Ala45]-IGF-1.

The only successfully prepared IGF-2 analogue in this series, [Gln45]-IGF-2, had reduced binding potency for IGF-1R compared to that of native IGF-2 ([Table 1](#)) but activated IGF-1R like native IGF-2 did ([Figure 3A](#)).

Relative EC₅₀ values of IGF-1R stimulation by the selected analogues were in good general agreement with their relative ability to stimulate this receptor performed at a set ligand concentration of 10 nM ([Table S2](#)). Hence, it appeared that the autophosphorylation abilities of hormones determined at their 10 nM concentrations were good representations of their properties, and as this approach substantially improved the time

Table 1. IGF-1R Binding Affinities of Native Hormones and Analogues Reported in This Work^a

analogue	K _d	SD (nM)	(n) for the human IGF-1R in mouse fibroblasts	relative binding affinity for human IGF-1R (%) relative to that of human IGF-1
human IGF-1	0.24	0.05 ^b	(5)	100
	0.12	0.01 ^c	(5)	100
	0.34	0.12 ^d	(4)	100
	0.16	0.06 ^e	(3)	100
	0.25	0.03 ^f	(4)	100
first series of IGF-1 analogues				
Gly-1-IGF-1	0.25	0.02 ^b	(3)	96
[His45]-IGF-1	0.39	0.11 ^c	(4)**	31
[Asn45]-IGF-1	0.20	0.07 ^c	(3)*	60
[Ala45]-IGF-1	0.41	0.27 ^c	(4)*	29
[His46]-IGF-1	0.18	0.04 ^b	(4)	133
[Gln46]-IGF-1	0.18	0.01 ^b	(3)	133
[Ala46]-IGF-1	0.26	0.11 ^b	(3)	92
[Asn45,Gln46]-IGF-1	0.70	0.28 ^d	(4)	49
second series of IGF-1 analogues				
[His49]-IGF-1	0.50	0.23 ^d	(4)	68
[His46,His49]-IGF-1	0.29	0.17 ^e	(4)	55
[Gln46,His49]-IGF-1	0.30	0.09 ^e	(3)	53
[Asn45,Gln46,His49]-IGF-1	0.92	0.04 ^d	(3)**	37
human IGF-2	2.3	1.2 ^f	(3)*	10.9
first series of IGF-2 analogues				
[Gln45]-IGF-2	2.9	0.3 ^c	(3)*	5.5
second series of IGF-2 analogues				
[His48]-IGF-2	0.88	0.23 ^c	(3)	18.2
[Gln45,His48]-IGF-2	0.89	0.05 ^c	(3)	18.0

^aThe values of K_d and relative binding affinities [relative receptor binding affinity defined as (K_d of human IGF-1)/(K_d of analogue) × 100] of human IGF-1, IGF-2, and analogues were determined for human IGF-1R in mouse fibroblasts. All IGF-1 analogues have an extra glycine residue at the N-terminus (Gly-1). n is the number of replicates. Asterisks indicate that binding of a particular ligand to IGF-1R differs significantly (*p < 0.05; ***p < 0.001) from the effect of IGF-1 in the case of IGF-1 analogues or differs significantly from the effect of IGF-2 in the case of IGF-2 analogues. Binding of native IGF-2 is related to that of human IGF-1. ^bRelative to the human IGF-1 K_d value of 0.24. ^cRelative to the human IGF-1 K_d value of 0.12. ^dRelative to the human IGF-1 K_d value of 0.34. ^eRelative to the human IGF-1 K_d value of 0.16. ^fRelative to the human IGF-1 K_d value of 0.25.

and material economy of this extensive methodology, the receptor activation abilities of the rest of the analogues were measured at this set ligand concentration only. In general, no major discrepancies between the IGF-1R binding and activation properties of the analogues mutated at positions 45 and 46 of IGF-1 and [Gln45]-IGF-2 were observed ([Table 1](#) and [Figure 3A](#)). Some analogues, e.g., [His46]-IGF-1 and [Gln46]-IGF-1, activated IGF-1R slightly less strongly than human IGF-1 did, but their apparent higher binding affinities for this receptor were not statistically significant. In contrast, [His45]-IGF-1, [Asn45]-IGF-1, and [Ala45]-IGF-1 analogues activated IGF-1R like human IGF-1 did, but their binding affinities were significantly reduced. Therefore, although some minor discrepancies could be observed here, any clear and major antagonism, or receptor overstimulation, was not detected.

IR-A Binding and Activation Properties of IGF-1 Analogues Modified at Sites 45 and 46, and the

Table 2. IR-A Receptor Binding Affinities of Native Hormones and Analogues Reported in This Work^a

analogue	K_d	SE (nM)	(n) for human IR-A in IM-9 lymphocytes	relative binding affinity for human IR-A (%) relative to that of human insulin
human insulin	0.25	0.05 ^b	(5)	100
	0.27	0.02 ^c	(5)	100
	0.18	0.01 ^d	(4)	100
	0.32	0.09 ^e	(4)	100
	0.30	0.13 ^f	(5)	100
human IGF-1	23.7	11.5 ^b	(3)***	1.1
first series of IGF-1 analogues				
Gly-1-IGF-1	35.6	11.9 ^b	(3)	0.7
[His45]-IGF-1	20.1	7.8 ^c	(4)	1.3
[Asn45]-IGF-1	19.3	9.6 ^c	(4)	1.4
[Ala45]-IGF-1	17.6	9.7 ^c	(3)	1.5
[His46]-IGF-1	6.6	1.2 ^d	(3)	2.7
[Gln46]-IGF-1	18.1	3.6 ^d	(3)	1.0
[Ala46]-IGF-1	14.0	1.9 ^d	(3)	1.3
[Asn45,Gln46]-IGF-1	17.5	8.4 ^e	(3)	1.8
second series of IGF-1 analogues				
[His49]-IGF-1	6.7	2.4 ^e	(3)*	4.8
[His46,His49]-IGF-1	3.4	1.7 ^e	(3)*	9.4
[Gln46,His49]-IGF-1	7.5	4.1 ^e	(3)*	4.3
[Asn45,Gln46,His49]-IGF-1	5.5	2.5 ^f	(4)***	5.5
human IGF-2	2.9	0.2 ^b	(3)***	8.6
first series of IGF-2 analogues				
[Gln45]-IGF-2	1.6	0.3 ^e	(3)**	20
second series of IGF-2 analogues				
[His48]-IGF-2	0.54	0.13 ^e	(3)***	59.3
[Gln45,His48]-IGF-2	0.65	0.12 ^f	(3)***	46.2

^aThe values of K_d and relative binding affinities [relative receptor binding affinity defined as $(K_d \text{ of human insulin}) / (K_d \text{ of analogue}) \times 100$] of human insulin, IGF-1, IGF-2, and analogues were determined for human IR-A in human IM-9 lymphocytes. All IGF-1 analogues have an extra glycine residue at the N-terminus (Gly-1). *n* is the number of replicates. Asterisks indicate that binding of a particular ligand to IR-A differs significantly (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$) from the effect of IGF-1 in the case of IGF-1 analogues or differs significantly from the effect of IGF-2 in the case of IGF-2 analogues. Binding of native IGF-2 is related to that of human insulin. ^bRelative to the human insulin K_d value of 0.25 0.05 ($n = 5$). ^cRelative to the human insulin K_d value of 0.27 0.02 ($n = 5$). ^dRelative to the human insulin K_d value of 0.18 0.01 ($n = 4$). ^eRelative to the human insulin K_d value of 0.32 0.09 ($n = 4$). ^fRelative to the human insulin K_d value of 0.30 0.13 ($n = 5$).

[Gln45]-IGF-2 Analogue. The IGF-1 analogues of the first series have binding affinities for IR-A similar to that of native IGF-1 (Table 2), and their IR-A activation properties (Figure 3B) are again in general agreement with the properties of native IGF-1.

The [Gln45]-IGF-2 analogue binds IR-A significantly more strongly ($\leq 20\%$ of the binding affinity of native human insulin) than the native IGF-2 that has only 8% of the binding affinity of human insulin. However, the IR-A enhanced affinity of this analogue was not fully translated into its activation potency that is similar to the activation potency of native IGF-2.

Design of the Second Series of IGF-1 and IGF-2 Analogues. The Thr8His substitution in insulin increases the potency for IR-A;^{41–46} hence, it was also used by Whittaker et al.³⁷ to increase the level of IR binding of A4/A5-modified

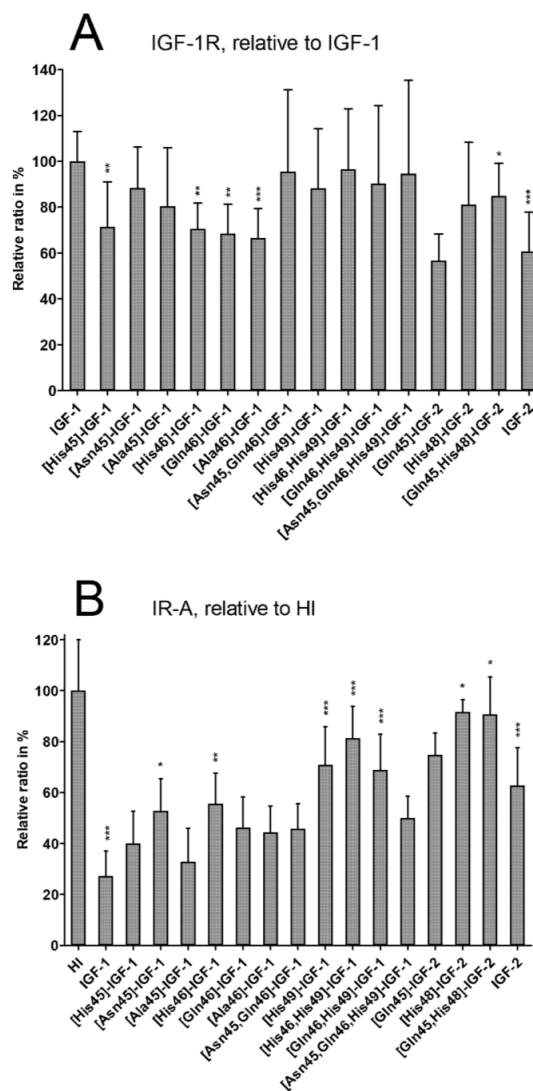


Figure 3. Relative abilities to activate (A) IGF-1R and (B) IR-A of human insulin (HI), human IGF-1, human IGF-2, and IGF-1 and IGF-2 analogues. All IGF-1 analogues contain a glycine residue at position -1. Relative abilities to activate receptors were determined with 10 nM ligands after a 10 min stimulation. Mean \pm SD values were calculated from four independent experiments ($n = 4$). In panel A, the experimental values are related to the biological activity of human IGF-1. In panel B, the experimental values are the biological activity of human insulin (HI). Asterisks indicate that induction of autophosphorylation of a particular receptor induced by a ligand differs significantly (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$) from the effect of IGF-1 in the case of IGF-1 analogues or differs significantly from the effect of IGF-2 in the case of IGF-2 analogues. In panel A, the significance of the effect of native IGF-2 (asterisks) is related to human IGF-1 and in panel B to human insulin. In panel B, the significance of the effect of native IGF-1 (asterisks) is related to human insulin.

insulins, without eliminating their antagonistic properties. Therefore, we probed a similar strategy for the IGF-1 and

IGF-2 analogues, which both have phenylalanine at the insulin A8 equivalent 49 and 48 sites, respectively (Figure 1).

First, the [His49]-IGF-1 analogue was made to investigate the effect of this single mutation. Subsequently, it was combined with His or Gln single mutations at site 46 and with Asn45/Gln46 double mutations, as well.

In the case of IGF-2, the single mutation [His48]-IGF-2 analogue was made, which was extended for Gln45 mutation, as well, as it was here the only successful substitution of IGF-2 in the first series.

IGF-1R Binding and Activation Properties of the Second Series of IGF-1 and IGF-2 Analogues. All new IGF-1 mutations have fairly minor, or no significant, effects on IGF-1R binding affinities and activation capabilities, with only the [Asn45,Gln46,His49]-IGF-1 triple mutant having significantly less affinity (37%) for the IGF-1R than native IGF-1 has but with native IGF-1-like autophosphorylation activation ability (Table 1 and Figure 3A).

A similar trend was observed for both new IGF-2 analogues, [His48]-IGF-2 and [Gln45,His48]-IGF-2, the IGF-1R binding and activation abilities of which were similar to those of native IGF-2.

IR-A Binding and Activation Properties of the Second Series of IGF-1 and IGF-2 Analogues. The IR-A-related properties of the analogues contrast with their IGF-1R affinities and binding effects. It seems that the presence of His49 strongly enhances (4–9-fold) the IR-A binding affinity of new IGF-1 analogues (Table 2), in comparison to that of native IGF-1. Moreover, His49-containing IGF-1 analogues do not show any IR-A antagonism, as their capabilities to activate this receptor are superior, or similar, to those of native IGF-1 (Figure 3B).

IR-A binding affinities of two new IGF-2 analogues carrying a His48 mutation are also very (5–7-fold) enhanced, in comparison with those of native IGF-2. Remarkably, both of these analogues bind IR-A with subnanomolar affinities, which make them half-equipotent with human insulin with respect to this IR isoform (Table 2). Moreover, the IR-A activation abilities of these two IGF-2 analogues are also very high, with a potency similar to that of human insulin (Figure 3B).

In general, the levels of activation of IR-A by insulin, IGF-1, IGF-2, and analogues determined at 10 nM (Figure 3B) correspond well to their respective binding affinities for this receptor (Table 2), and they are in a good agreement with the levels of IR-A autophosphorylation induced by 10 nM insulin, IGF-1, and IGF-2 in dose–response curves reported recently by Andersen et al.⁵³

IR-B Binding Properties of [His49]-IGF-1 and [Gln45,His48]-IGF-2 Analogues. The outstanding enhancement of IR-A binding affinities exhibited by the His49/His48 IGF-1/2 mutants prompted their characterization toward the IR-B isoform, as well. For this purpose, we tested two representative analogues, [His49]-IGF-1 and [Gln45,His48]-IGF-2, for their binding affinity for IR-B (Table 3). Most interestingly, we found that the IR-B binding affinity of [Gln45,His48]-IGF-2 is ~8 times higher than that of human IGF-2 and that the binding of [His49]-IGF-1 to this IR isoform is ~3 times stronger than that of native IGF-1. These mutations did not considerably affect the IR isoform specificity of the analogues because their IR-A binding enhancing effects were similar (5 times for [Gln45,His48]-IGF-2 and ~4 times for [His49]-IGF-1).

Table 3. IR-B Receptor Binding Affinities of Native Hormones and [His49]-IGF-1 and [Gln45,His48]-IGF-2 Analogues Reported in This Work^a

analogue	K_d	SE (nM)	(n) for human IR-B in mouse fibroblasts	relative binding affinity for human IR-B (%) relative to that of human insulin
human insulin	0.50	0.31	(5)	100 62
human IGF-1	224	16	(4) ^{b***}	0.22 0.02
[His49]-IGF-1	72.3	12.0	(3) ^{***}	0.69 0.11
human IGF-2	35.5	5.6	(4) ^{b***}	1.4 0.2
[Gln45,His48]-IGF-2	4.3	1.7	(4) ^{**}	11.6 4.6

^aThe values of K_d and relative binding affinities [relative receptor binding affinity defined as (K_d of human insulin)/(K_d of analogue) × 100] of human insulin, IGF-1, IGF-2, and analogues were determined for human IR-B in mouse fibroblasts. *n* is the number of replicates. Asterisks indicate that the binding a particular ligand to IR-B differs significantly (***p* < 0.01; ****p* < 0.001) from the effect of IGF-1 in the case of an IGF-1 analogue or differs significantly from the effect of IGF-2 in the case of an IGF-2 analogue. The binding of native IGF-1 and IGF-2 is relative to that of human insulin. ^bFrom ref 52.

DISCUSSION

All 11 planned IGF-1 analogues were successfully produced, but only three IGF-2 analogues ([Gln45]-IGF-2, [Gln45,His49]-IGF-2, and [His49]-IGF-2) were made, as the other analogues of this hormone formed insoluble precipitates during the folding steps of their purification. As Whittaker et al.³⁷ successfully prepared HisA4, AlaA4, and HisA5 insulins, the severe aggregation of some IGF-2 analogues indicates different folding mechanisms of these hormones, underlining the importance of IGF-2 Glu44/45 for its efficient assembly. However, a more extensive mutagenesis, especially of the IGF-2 Glu44 site, is needed for an unambiguous confirmation of the folding-related importance of these side chains.

In general, the majority of mutations performed here have a rather minor impact on the IGF-1R binding affinities of IGF-1 and IGF-2 analogues (Table 1). However, different trends may be observed for the His mutation at position 48 of IGF-2 that enhances its IGF-1R affinity, while the His mutation at IGF-1 equivalent site 49 has an opposite effect. This may indicate that the natures of interactions of IGF-1 and IGF-2 with IGF-1R are different, and specific, for IGF-1 and IGF-2 at their sites 49 and 48, respectively; i.e., the equivalent amino acids at these sites do not interact with the IGF-1R in the same fashion.

Interestingly, the mutations of sites 49 and 48 in IGF-1 and IGF-2, respectively, yielded analogues of these hormones with much more significant, and interesting, changes in their IR-A binding affinities (Table 2) than in their IGF-1R binding/activation properties. Here, both IGF-1 and IGF-2 analogues with His at sites 49 and 48, respectively, are much better IR-A binders than their native forms. The data presented here indicate that this effect can be attributed mainly to the isolated impact of His49, or His48, as the simultaneous mutations at these sites and of residues 45 and 46 in IGF-1 (or residue 45 in IGF-2) do not generate any additional, significant positive properties toward IR-A.

It has been shown that the ThrA8His mutation doubles or triples the insulin IR-A binding affinity,^{41–45} and Whittaker et al.³⁷ used such a mutation to restore IR binding affinity of the less active A4 and A5 insulin mutants. Here, similar mutations enhanced 4–9-fold the IR-A binding affinities of the IGF-1 analogues and 5–7-fold the binding affinities of the IGF-2

analogues, compared to the affinities of the native hormones. The positive effect of His49 and His48 mutations on IR-A binding is similar for both IGF-1 and IGF-2, and to the best of our knowledge, these analogues are the strongest IGF-1-like and IGF-2-like binders of the IR-A receptor isoform thus far reported. In particular, the high IR-A binding affinity of [His48]-IGF-2 is remarkable, as only this single mutation was sufficient to generate an analogue with ~50% insulin-like affinity for the IR-A. Moreover, we found that the IR-A binding affinity enhancing effect of His49 and His48 is manifested with the IR-B, as well, because the [His49]-IGF-1 analogue is an ~3-fold stronger IR-B binder than native IGF-1 is and, remarkably, [Gln45,His48]-IGF-2 binds IR-B ~8-fold more strongly than native IGF-2 does, having almost 12% of the binding affinity of human insulin. Therefore, it seems that [His49]-IGF-1 and [Gln45,His48]-IGF-2 analogues are the strongest reported IGF-1-like and IGF-2-like IR-B binders. The exceptional binding “promiscuity” of [His48]-IGF-2 toward IR-A and IR-B is outstanding and reveals how easily IGF-2 can be converted into a high-affinity ligand for IR receptors. This “ubiquitous” hormonal property of IGF-2 may evoke questions about its evolutionary origins, and it could be hypothesized that IGF-2 resembles the hypothetical evolutionary hormonal ancestor of the insulin/IGF axis more closely than insulin or IGF-1 does.

It is generally accepted that insulin and IGFs interact with their receptors through two main binding sites, sites 1 and 2 in the hormones, and sites 1' and 2', respectively, in the receptors.⁵⁴ The nature of the interactions of site 1 in insulin and IGF-1 with site 1' in IR-A and IGF-1R is relatively well characterized in the crystal structures of their complexes (Figure 2). However, structural details about hormone site 2–receptor site 2' interactions are still missing, and insulin amino acids, which determine its site 2, ThrA8, IleA10, SerA12, LeuA13, GluA17, HisB10, GluB13, and LeuA17, have been suggested through extensive mutagenesis studies.⁵⁴

The different effects of His mutations in IGFs on binding affinities for IR-A, IR-B, and IGF-1R could mean that amino acids at positions 49 (IGF-1) and 48 (IGF-2) are engaged with significantly different protein environments in complexes with IR and IGF-1R. Although it is expected that Phe49 in IGF-1, Phe48 in IGF-2, and ThrA8 in insulin (all at site 2 of the hormones) interact with 2' sites of the receptors, the possibility that the increase in the level of IR-A binding of [His48]-IGF-2 and [His49]-IGF-1 analogues may also result from an enhancement of some contacts with elements of IR site 1' cannot be excluded (Figure 2). For example, His49 and His48 of IGFs could form double direct, or water-mediated, hydrogen bonds with Asp707 in the IR α -CT segment and with Asn694 in IGF-1R, as well (Figure 2A), while insulin native ThrA8 is too short for such interactions. It is also possible that the native Phe49 and Phe48 in IGFs can contribute only some weak van der Waals intramolecular contacts with the IGF-1 Cys48–Cys6 disulfide bond and Val44 (IGF-2 Cys47–Cys9 and Val43). Hence, His49 and His48 mutations may provide much stronger directional contacts with both receptor sites (1' and 2'), which seem to be more favorable for IR-A rather than for IGF-1R.

In general, our new IGF-1 and IGF-2 analogues did not show any considerable reversed trends between their affinities and receptor activation abilities (Tables 1 and 2, Figure 3A,B, and Figure S5). Therefore, the previously observed significant disproportionate binding and activation of insulin analogues mutated at A4 and A8 (ref 37) are probably specific and limited for insulin–IR interactions. It is possible that a different

receptor activation behavior of insulin and IGF-1 (IGF-2) analogues mutated at positions A4 and A8 and positions 45, 46, and 49 (positions 45 and 48), respectively, may result from different interactions with specific amino acids at 2' sites of the IR that are currently not yet determined (see the preceding discussion). Such dissimilar natures of hormone–receptor interfaces may lead to non-equivalent receptor binding mechanisms (as previously pertinently mentioned by Sohma et al. in ref 55) and their subsequent different impacts on signal transduction through the receptors and, ultimately, the activation of their tyrosine kinases.

As indicated above, some trends observed in the analogues described here can be also corroborated by the phylogeny of insulin-related hormones. For example, the insulin ThrA8-containing A8–A10 region is considered as a hypervariable part of this hormone because of its significant sequence differences in mammals.⁵⁶ However, the A8 site is the most invariable amino acid in the A8–A10 triad. For example, the well IR-A tolerated (87% binding affinity of HI) AlaA8 occurs in cattle, sheep, and goat insulin.⁵⁷ Although there is no mammalian insulin with His at site A8, this amino acid is frequently present at that position in fish, frog, and bird insulins.⁵⁶ Also, HisA8 was proposed to be responsible for an ~5-fold higher binding affinity of chicken insulin for IR in human lymphocytes.⁴¹ Herring et al.⁴⁶ proposed that it is possible that a “lower-affinity” ThrA8 site in mammals emerged from the evolutionary optimization of the insulin receptor kinetics, which requires a decrease in insulin binding affinity in mammals that is dictated by a different route of insulin delivery. In more ancient vertebrates like birds, insulin is secreted differently, by the kidney, which leads to rapid metabolic changes.⁴⁶ However, in mammals, the pancreas-to-liver portal vein-mediated pathway is the primary direction of insulin. In humans, the liver is the main glycemic-response organ with >90% “metabolic” IR-B isoform and where insulin inhibits gluconeogenesis and glycogenolysis before reaching the peripheral tissues. It may be postulated then that to ensure an optimum liver:body distribution of insulin its IR binding affinity needs to be reduced in mammals, to achieve a more systemic glycemic response.⁴⁶

All known naturally occurring variants of IGF-1 and IGF-2 (including avian) maintain Phe at their insulin ThrA8-corresponding sites 49 and 48, respectively. Therefore, it may be assumed that the IR high-affinity inducing HisA8 mutation in insulin selected upon evolution was required for only rapid and immediate metabolic effects of insulin and was not needed for slower and long-lasting growth effects of IGFs. However, a high sensitivity of both IGFs to acquisition of His48 and His49 mutations may reflect their common evolutionary origins with insulin.

It will be interesting to further explore the findings of this report by testing specific amino acid substitutions at the A8 site of insulin, to maintain its potent IR binding but to decrease its affinity for IGF-1R. Such analogues could be useful for the safer treatment of diabetes, as some currently administered insulin derivatives (e.g., glargine) show higher IGF-1R affinity and hence, potentially, an increased risk of cancer.^{58–62}

In summary, as we showed that respective positions in insulin and IGFs lead to non-equivalent IR-A and IGF-1R binding mechanisms of these hormones, this work opens new directions for the rational engineering of the hormonal components of the insulin/IGF system.

■ CONCLUSIONS

We systematically investigated the receptor binding and receptor activation properties of IGF-1 and IGF-2 analogues modified at positions 45, 46, and 49 (IGF-1) and at positions 44, 45, and 48 (IGF-2). These modifications did not significantly affect the IGF-1R binding of these hormones. However, analogues with the Phe49His mutation in IGF-1 and the Phe48His mutation in IGF-2 have remarkably enhanced binding affinities for both IR-A and IR-B isoforms of the insulin receptor. Here, IGF-1 analogues with His at position 49 possess approximately 5–9% of IR-A, and [His49]-IGF-1 possess ~0.7% of the IR-B binding affinity of human insulin. Moreover, IGF-2 analogues with His at position 48 are approximately half-equipotent to human insulin in binding to IR-A, and [Gln45,His48]-IGF-2 has almost 12% of the insulin binding affinity for this “metabolic” isoform. The binding affinities of all analogues are in general proportionate to their abilities to activate IR and IGF-1R without any important discrepancies. This study revealed that IR and IGF-1R can contain specific sites, probably parts of so-called receptors’ sites 2’, which can interact differently with insulin and with IGFs. These findings shed light on new, possible directions of rational engineering of insulin and IGFs toward more selective and receptor-specific analogues with medical applications.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.biochem.7b01260](https://doi.org/10.1021/acs.biochem.7b01260).

Summary of mutagenesis strategies used for cloning of IGF-1 and IGF-2 analogues (Table S1), binding curves of native hormones and analogues with IGF-1R, IR-A, and IR-B receptors (Figures S1–S3, respectively), representative immunoblot for determination of receptor activation abilities (Figure S4), and dose–response curves for determination of EC₅₀ values of IGF-1R activation abilities of selected analogues (Figure S5) and their respective EC₅₀ values (Table S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*Telephone: +420 220183 441. E-mail: jiracek@uochb.cas.cz.

*E-mail: zakova@uochb.cas.cz.

ORCID

Andrzej M. Brzozowski: [0000-0001-7426-8948](https://orcid.org/0000-0001-7426-8948)

Jiří Jiráček: [0000-0003-3848-2773](https://orcid.org/0000-0003-3848-2773)

Present Address

^{||}J.R.: Department of Biochemistry, Faculty of Chemistry, University of Belgrade, Studentski trg 12-16, Belgrade, Serbia.

Funding

This work was supported by Czech Science Foundation Grant 15-19018S (to L.Z.), by Medical Research Council Grant MR/K000179/1 (to A.M.B.), and partly also by Medical Research Council Grant MR/R009066/1 (to A.M.B.). Institutional support was provided by Projects RVO 61388963 (for the Institute of Organic Chemistry and Biochemistry) and 68378050 (for the Institute of Molecular Genetics) of the Czech Academy of Sciences.

Notes

The authors declare no competing financial interest.

■ REFERENCES

- Chan, S. J., Nagamatsu, S., Cao, Q. P., and Steiner, D. F. (1992) Structure and evolution of Insulin and Insulin-Like Growth-Factors in Chordates. *Prog. Brain Res.* 92, 15–24.
- Chan, S. J., and Steiner, D. F. (2000) Insulin through the ages: phylogeny of a growth promoting and metabolic regulatory hormone. *Am. Zool.* 40, 213–222.
- McRory, J. E., and Sherwood, N. M. (1997) Ancient divergence of insulin and insulin-like growth factor. *DNA Cell Biol.* 16, 939–949.
- LeRoith, D., Kavsan, V. M., Koval, A. P., and Roberts, C. T. (1993) Phylogeny of the insulin-like growth-factors (IGFs) and receptors - a molecular approach. *Mol. Reprod. Dev.* 35, 332–338.
- Denley, A., Cosgrove, L. J., Booker, G. W., Wallace, J. C., and Forbes, B. E. (2005) Molecular interactions of the IGF system. *Cytokine Growth Factor Rev.* 16, 421–439.
- Mayer, J. P., Zhang, F., and DiMarchi, R. D. (2007) Insulin structure and function. *Biopolymers* 88, 687–713.
- Stewart, C. E. H., and Rotwein, P. (1996) Growth, differentiation, and survival: multiple physiological functions for insulin-like growth factors. *Physiol. Rev.* 76, 1005–1026.
- Clemmons, D. R. (2009) Role of IGF-I in skeletal muscle mass maintenance. *Trends Endocrinol. Metab.* 20, 349–356.
- Holly, J. M. P., and Perks, C. M. (2012) Insulin-like growth factor physiology. What we have learned from human studies. *Endocrinol. Metab. Clin. North Am.* 41, 249–263.
- Dynkevich, Y., Rother, K. I., Whitford, I., Qureshi, S., Galiveeti, S., Szulc, A. L., Danoff, A., Breen, T. L., Kaviani, N., Shanik, M. H., LeRoith, D., Vigneri, R., Koch, C. A., and Roth, J. (2013) Tumors, IGF-2, and hypoglycemia: insights from the clinic, the laboratory, and the historical archive. *Endocr. Rev.* 34, 798–826.
- Alberini, C. M., and Chen, D. Y. (2012) Memory enhancement: consolidation, reconsolidation and insulin-like growth factor 2. *Trends Neurosci.* 35, 274–283.
- Chen, D. Y., Stern, S. A., Garcia-Osta, A., Saunier-Rebori, B., Pollonini, G., Bambah-Mukku, D., Blitzer, R. D., and Alberini, C. M. (2011) A critical role for IGF-II in memory consolidation and enhancement. *Nature* 469, 491–497.
- Lawrence, M. C., McKern, N. M., and Ward, C. W. (2007) Insulin receptor structure and its implications for the IGF-1 receptor. *Curr. Opin. Struct. Biol.* 17, 699–705.
- De Meyts, P. (2004) Insulin and its receptor: structure, function and evolution. *BioEssays* 26, 1351–1362.
- De Meyts, P., and Whittaker, J. (2002) Structure-Function Relationships of Insulin and Insulin-Like Growth Factor-I Receptor Binding. In *Insulin and Related Proteins: Structure to Function and Pharmacology* (Dieken, M. L., Federwisch, M., and De Meyts, P., Eds.) pp 131–149, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Belfiore, A., Frasca, F., Pandini, G., Sciacca, L., and Vigneri, R. (2009) Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr. Rev.* 30, 586–623.
- Belfiore, A., Malaguarnera, R., Vella, V., Lawrence, M. C., Sciacca, L., Frasca, F., Morrione, A., and Vigneri, R. (2017) Insulin receptor isoforms in physiology and disease: an updated view. *Endocr. Rev.* 38, 379–431.
- Vienberg, S. G., Bouman, S. D., Sorensen, H., Stidsen, C. E., Kjeldsen, T., Glendorf, T., Sorensen, A. R., Olsen, G. S., Andersen, B., and Nishimura, E. (2011) Receptor-isoform-selective insulin analogues give tissue-preferential effects. *Biochem. J.* 440, 301–308.
- Jiracek, J., and Zakova, L. (2017) Structural perspectives of insulin receptor isoform-selective insulin analogs. *Front. Endocrinol.* 8, 167.
- Siddle, K. (2012) Molecular basis of signaling specificity of insulin and IGF receptors: neglected corners and recent advances. *Front. Endocrinol.* 3, 34.
- Siddle, K. (2011) Signalling by insulin and IGF receptors: supporting acts and new players. *J. Mol. Endocrinol.* 47, R1–R10.

- (22) Boucher, J., Kleinridders, A., and Kahn, C. R. (2014) Insulin receptor signaling in normal and insulin-resistant states. *Cold Spring Harbor Perspect. Biol.* 6, a009191.
- (23) Hers, I., Vincent, E. E., and Tavare, J. M. (2011) Akt signalling in health and disease. *Cell. Signalling* 23, 1515–1527.
- (24) Bedinger, D. H., and Adams, S. H. (2015) Metabolic, anabolic, and mitogenic insulin responses: A tissue-specific perspective for insulin receptor activators. *Mol. Cell. Endocrinol.* 415, 143–156.
- (25) Sacco, A., Morcavallo, A., Pandini, G., Vigneri, R., and Belfiore, A. (2009) Differential signaling activation by insulin and insulin-like growth factors I and II upon binding to insulin receptor isoform A. *Endocrinology* 150, 3594–3602.
- (26) Slaaby, R., Schaffer, L., Lautrup-Larsen, I., Andersen, A. S., Shaw, A. C., Mathiasen, I. S., and Brandt, J. (2006) Hybrid receptors formed by insulin receptor (IR) and insulin-like growth factor I receptor (IGF-IR) have low insulin and high IGF-1 affinity irrespective of the IR splice variant. *J. Biol. Chem.* 281, 25869–25874.
- (27) Slaaby, R. (2015) Specific insulin/IGF1 hybrid receptor activation assay reveals IGF1 as a more potent ligand than insulin. *Sci. Rep.* 5, 7911.
- (28) Clemmons, D. R. (2016) Role of IGF binding proteins in regulating metabolism. *Trends Endocrinol. Metab.* 27, 375–391.
- (29) Firth, S. M., and Baxter, R. C. (2002) Cellular actions of the insulin-like growth factor binding proteins. *Endocr. Rev.* 23, 824–854.
- (30) Kornfeld, S. (1992) Structure and function of the mannose 6-phosphate insulin-like growth factor-II receptors. *Annu. Rev. Biochem.* 61, 307–330.
- (31) Belfiore, A., and Malaguarnera, R. (2011) Insulin receptor and cancer. *Endocr.-Relat. Cancer* 18, R125–R147.
- (32) Gallagher, E. J., and LeRoith, D. (2011) Minireview: IGF, insulin, and cancer. *Endocrinology* 152, 2546–2551.
- (33) Tao, Y., Pinzi, V., Bourhis, J., and Deutsch, E. (2007) Mechanisms of disease: signaling of the insulin-like growth factor I receptor pathway-therapeutic perspectives in cancer. *Nat. Clin. Pract. Oncol.* 4, 591–602.
- (34) Baserga, R. (2013) The decline and fall of the IGF-1 receptor. *J. Cell. Physiol.* 228, 675–679.
- (35) Schaffer, L., Brissette, R. E., Spetzler, J. C., Pillutla, R. C., Ostergaard, S., Lennick, M., Brandt, J., Fletcher, P. W., Danielsen, G. M., Hsiao, K. C., Andersen, A. S., Dedova, O., Ribbel, U., Hoeg-Jensen, T., Hansen, P. H., Blume, A. J., Markussen, J., and Goldstein, N. I. (2003) Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4435–4439.
- (36) Pillutla, R., Brissette, R., Blume, A. J., Schaffer, L., Brandt, J., Goldstein, N. I., Spetzler, J., Ostergaard, S., and Hansen, P. H. (2011) Insulin and IGF-1 Receptor Agonists and Antagonists. U.S. Patent 2011/0124556 A1.
- (37) Whittaker, J., Whittaker, L. J., Roberts, C. T., Phillips, N. B., Ismail-Beigi, F., Lawrence, M. C., and Weiss, M. A. (2012) α -Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase. *Proc. Natl. Acad. Sci. U. S. A.* 109, 11166–11171.
- (38) Menting, J. G., Whittaker, J., Margetts, M. B., Whittaker, L. J., Kong, G. K. W., Smith, B. J., Watson, C. J., Zakova, L., Kletvikova, E., Jiracek, J., Chan, S. J., Steiner, D. F., Dodson, G. G., Brzozowski, A. M., Weiss, M. A., Ward, C. W., and Lawrence, M. C. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* 493, 241–245.
- (39) Menting, J. G., Lawrence, C. F., Kong, G. K. W., Margetts, M. B., Ward, C. W., and Lawrence, M. C. (2015) Structural congruency of ligand binding to the insulin and insulin/type 1 insulin-like growth factor hybrid receptors. *Structure* 23, 1271–1282.
- (40) Hexnerova, R., Krizkova, K., Fabry, M., Sieglöva, I., Kedrova, K., Collinsova, M., Ullrichova, P., Srb, P., Williams, C., Crump, M. P., Tosner, Z., Jiracek, J., Veverka, V., and Zakova, L. (2016) Probing receptor specificity by sampling the conformational space of the insulin-like growth factor II C-domain. *J. Biol. Chem.* 291, 21234–21245.
- (41) Simon, J., Freychet, P., Rosselin, G., and Demeys, P. (1977) Enhanced binding affinity of chicken insulin in rat-liver membranes and human lymphocytes - relationship to kinetic properties of hormone-receptor interaction. *Endocrinology* 100, 115–121.
- (42) Kaarsholm, N. C., Norris, K., Jorgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sorensen, A. R., and Havelund, S. (1993) Engineering stability of the insulin monomer fold with application to structure-activity relationships. *Biochemistry* 32, 10773–10778.
- (43) Weiss, M. A., Hua, Q. X., Jia, W., Nakagawa, S. H., Chu, Y. C., Hu, S. Q., and Katsyannis, P. G. (2001) Activities of monomeric insulin analogs at position A8 are uncorrelated with their thermodynamic stabilities. *J. Biol. Chem.* 276, 40018–40024.
- (44) Weiss, M. A., Wan, Z., Zhao, M., Chu, Y. C., Nakagawa, S. H., Burke, G. T., Jia, W., Hellmich, R., and Katsyannis, P. G. (2002) Non-standard insulin design: structure-activity relationships at the periphery of the insulin receptor. *J. Mol. Biol.* 315, 103–111.
- (45) Wan, Z., Xu, B., Huang, K., Chu, Y. C., Li, B., Nakagawa, S. H., Qu, Y., Hu, S. Q., Katsyannis, P. G., and Weiss, M. A. (2004) Enhancing the activity of insulin at the receptor interface: crystal structure and photo-cross-linking of A8 analogues. *Biochemistry* 43, 16119–16133.
- (46) Herring, R., Jones, R. H., and Russell-Jones, D. L. (2014) Hepatoselectivity and the evolution of insulin. *Diabetes, Obes. Metab.* 16, 1–8.
- (47) Laible, M., and Boonrod, K. (2009) Homemade site directed mutagenesis of whole plasmids. *J. Visualized Exp.* 27, No. e1135.
- (48) Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* 2, 924–932.
- (49) Vikova, J., Collinsova, M., Kletvikova, E., Budesinsky, M., Kaplan, V., Zakova, L., Veverka, V., Hexnerova, R., Avino, R. J. T., Strakova, J., Selicharova, I., Vanek, V., Wright, D. W., Watson, C. J., Turkenburg, J. P., Brzozowski, A. M., and Jiracek, J. (2016) Rational steering of insulin binding specificity by intra-chain chemical crosslinking. *Sci. Rep.* 6, 19431.
- (50) Zakova, L., Kletvikova, E., Lepsik, M., Collinsova, M., Watson, C. J., Turkenburg, J. P., Jiracek, J., and Brzozowski, A. M. (2014) Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 70, 2765–2774.
- (51) Machackova, K., Collinsova, M., Chrudinova, M., Selicharova, I., Picha, J., Budesinsky, M., Vanek, V., Zakova, L., Brzozowski, A. M., and Jiracek, J. (2017) Insulin-like growth factor 1 analogs clicked in the C domain: chemical synthesis and biological activities. *J. Med. Chem.* 60, 10105–10117.
- (52) Krizkova, K., Chrudinova, M., Povalova, A., Selicharova, I., Collinsova, M., Vanek, V., Brzozowski, A. M., Jiracek, J., and Zakova, L. (2016) Insulin-insulin-like growth factors hybrids as molecular probes of hormone:receptor binding specificity. *Biochemistry* 55, 2903–2913.
- (53) Andersen, M., Nørgaard-Pedersen, D., Brandt, J., Pettersson, I., and Slaaby, R. (2017) IGF1 and IGF2 specificities to the two insulin receptor isoforms are determined by insulin receptor amino acid 718. *PLoS One* 12, e0178885.
- (54) De Meyts, P. (2015) Insulin/receptor binding: the last piece of the puzzle? *BioEssays* 37, 389–397.
- (55) Sohma, Y., Pentelute, B. L., Whittaker, J., Hua, Q. X., Whittaker, L. J., Weiss, M. A., and Kent, S. B. H. (2008) Comparative properties of insulin-like growth factor 1 (IGF-1) and [Gly7D-Ala]IGF-1 prepared by total chemical synthesis. *Angew. Chem., Int. Ed.* 47, 1102–1106.
- (56) Conlon, J. M. (2001) Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. *Peptides* 22, 1183–1193.
- (57) Kristensen, C., Kjeldsen, T., Wiberg, F. C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D. F., and Andersen, A. S. (1997) Alanine scanning mutagenesis of insulin. *J. Biol. Chem.* 272, 12978–12983.
- (58) Kurtzhals, P., Schaffer, L., Sorensen, A., Kristensen, C., Jonassen, I., Schmid, C., and Trub, T. (2000) Correlations of receptor binding

and metabolic and mitogenic potencies of insulin analogs designed for clinical use. *Diabetes* 49, 999–1005.

(59) Hemkens, L. G., Grouven, U., Bender, R., Gunster, C., Gutschmidt, S., Selke, G. W., and Sawicki, P. T. (2009) Risk of malignancies in patients with diabetes treated with human insulin or insulin analogues: a cohort study. *Diabetologia* 52, 1732–1744.

(60) Jonasson, J. M., Ljung, R., Talback, M., Haglund, B., Gudbjornsdottir, S., and Steineck, G. (2009) Insulin glargine use and short-term incidence of malignancies—a population-based follow-up study in Sweden. *Diabetologia* 52, 1745–1754.

(61) Colhoun, H. M. (2009) Use of insulin glargine and cancer incidence in Scotland: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetologia* 52, 1755–1765.

(62) Currie, C. J., Poole, C. D., and Gale, E. A. M. (2009) The influence of glucose-lowering therapies on cancer risk in type 2 diabetes. *Diabetologia* 52, 1766–1777.

Supporting Information

for

Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation

Kateřina Macháčková, Martina Chrudinová, Jelena Radosavljević, Pavlo Potalitsyn, Květoslava

Křížková, Milan Fábry, Irena Selicharová, Michaela Collinsová,

Andrzej M. Brzozowski, Lenka Žáková* and Jiří Jiráček*

Table of Contents

Table S1. Mutagenesis strategies for IGF-1 and IGF-2 analogs	Page S2
Figure S1. Binding curves of analogs to IGF-1R	Pages S3-S4
Figure S2. Binding curves of analogs to IR-A	Pages S5-S6
Figure S3. Binding curves of analogs to IR-B	Page S7
Figure S4. Representative immunoblot for determination of autophosphorylation activities of hormones.	Page S7
Figure S5. Representative dose response curves of the abilities of selected IGF-1 analogs to stimulate autophosphorylation of IGF-1R	Page S8
Table S2. Comparison of abilities of selected IGF-1 analogs to activate IGF-1R in a dose response manner (EC ₅₀ values) with their ability to activate IGF-1R at 10 nM concentration	Page S9
References	Page S9

Table S1. Mutagenesis strategies used for cloning of successfully expressed and produced IGF-1 and IGF-2 analogs.

Analog (expression construct)	Mutation(s)	Forward mutagenic primer (5'-3')	Reverse mutagenic primer (5'-3')	Template construct	Applied strategy
Gly-1-IGF-1	-				
[His45]-IGF-1	[D45H]	CCGGCATTGTGCACGGAATGCT GCTTTCGC	GCGAAAGCAGCATTTCGTGCA CAATGCCGG	Gly-1-IGF-1	SDM
[Asn45]-IGF-1	[D45N]	CCGGCATTGTGTTAACGAATGCT GCTTTCGC	GCGAAAGCAGCATTTCGTAA CAATGCCGG	Gly-1-IGF-1	SDM
[Ala45]-IGF-1	[D45A]	CCGGCATTGTGGCGGAATGCT GCTTTCGC	GCGAAAGCAGCATTCCGCCA CAATGCCGG	Gly-1-IGF-1	SDM
[His46]-IGF-1	[E46H]	CCGGCATTGTGGATCACTGCT GCTTTCGC	GCGAAAGCAGCAGTGTCCA CAATGCCGG	Gly-1-IGF-1	SDM
[Gln46]-IGF-1	[E46Q]	CCGGCATTGTGGATCAATGCT GCTTTCGC	GCGAAAGCAGCATTGTATCCA CAATGCCGG	Gly-1-IGF-1	SDM
[Ala46]-IGF-1	[E46A]	CCGGCATTGTGGATGCATGCT GCTTTCGC	GCGAAAGCAGCATGCAATCCA CAATGCCGG	Gly-1-IGF-1	SDM
[Asn45, Gln46]-IGF-1	[D45N+E46Q]	ACCGGCATTGTTAACCAAGTGC TGGTTTCGCAGC	GCTGCGAAAGCAGCACTGGT TAACAATGCCGGT	Asn45-IGF-1	OE
[His49]-IGF-1	[F49H]	GTGGATGAAATGCTGCCATCGC AGCTGCGATCTG	CAGATCGCAGCTGCCATGGC AGCATTCAATCCAC	Gly-1-IGF-1	OE
His46, His49]-IGF-1	[E46H+F49H]	GTGGATCACTGTGCCACCCGC AGCTGCGATCTG	CAGATCGCAGCTGCCGTGGC AGCAGTGTATCCAC	His46-IGF-1	SDM
[Gln46, His49]-IGF-1	[E46Q+F49H]	GTGGATCAATGCTGCCATCGC AGCTGCGATCTG	CAGATCGCAGCTGCCATGGC AGCATTGTATCCAC	Gln46-IGF-1	OE
[Asn45, Gln46, His49]-IGF-1	[D45N+E46Q+F49H]	ATTGTTAACCAAGTGTGCCATC GCAGCTGCGATCT	AGATCGCAGCTGCCATGGCA GCACTGGTTAACAAAT	Asn45-IGF-1	OE
IGF-2	-				
[Gln45]-IGF-2	[E45Q]	CGGGCATTGTGGAACAGTGC TGCTTTCGCAGC	GCTGCGAAAGCAGCACTGTT CCACAATGCCGG	IGF-2	SDM
[His48]-IGF-2	[F48H]	GTGGAAGAATGTGCCATCGC AGCTGCGATCTG	CAGATCGCAGCTGCCATGGC AGCATTCTCCAC	IGF-2	OE
[Gln45, His48]-IGF-2	[E45H+F48H]	CGGGCATTGTGGAACAGTGC TGCCATCGCAGC	GCTGCGATGGCAGCACTGTT CCACAATGCCGG	His48-IGF-2	SDM

OE - overlap-extension PCR (1), SDM - site-directed mutagenesis of the whole plasmid (2).

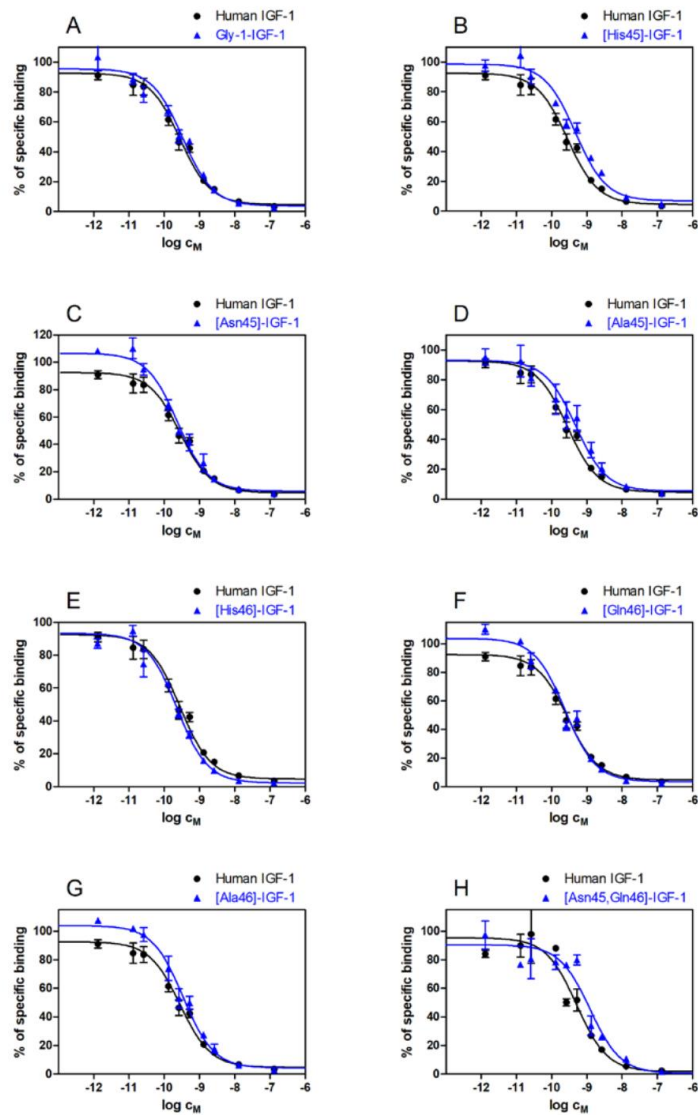


Figure S1. Inhibition of binding of human [125 I]moniodotyrosyl-IGF-1 to human IGF-1R in mouse embryonic fibroblasts by Gly-1-IGF-1 (A), [His45]-IGF-1 (B), [Asn45]-IGF-1 (C), [Ala45]-IGF-1 (D), [His46]-IGF-1 (E), [Gln46]-IGF-1 (F), [Ala46]-IGF-1 (G), [Asn45,Gln46]-IGF-1 (H) and by human IGF-1 (in all panels). All IGF-1 analogs have Gly at the position -1. The representative binding curves are shown.

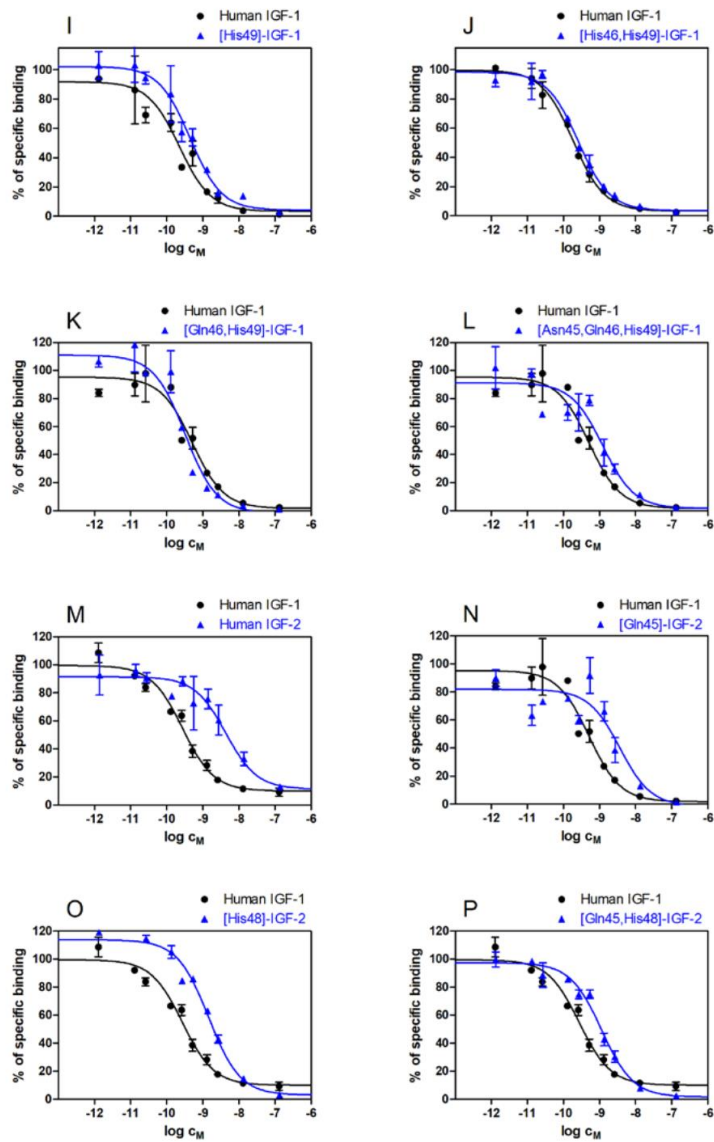


Figure S1 (continued). Inhibition of binding of human [¹²⁵I]monoiodotyrosyl-IGF-1 to human IGF-1R in mouse embryonic fibroblasts by [His49]-IGF-1 (I), [His46,His49]-IGF-1 (J), [Gln46,His49]-IGF-1 (K), [Asn45,Gln46,His49]-IGF-1 (L), human IGF-2 (M), [Gln45]-IGF-2 (N), [His48]-IGF-2 (O), [Gln45,His48]-IGF-2 (P) and by human IGF-1 (in all panels). All IGF-1 analogs have Gly at the position -1. The representative binding curves are shown.

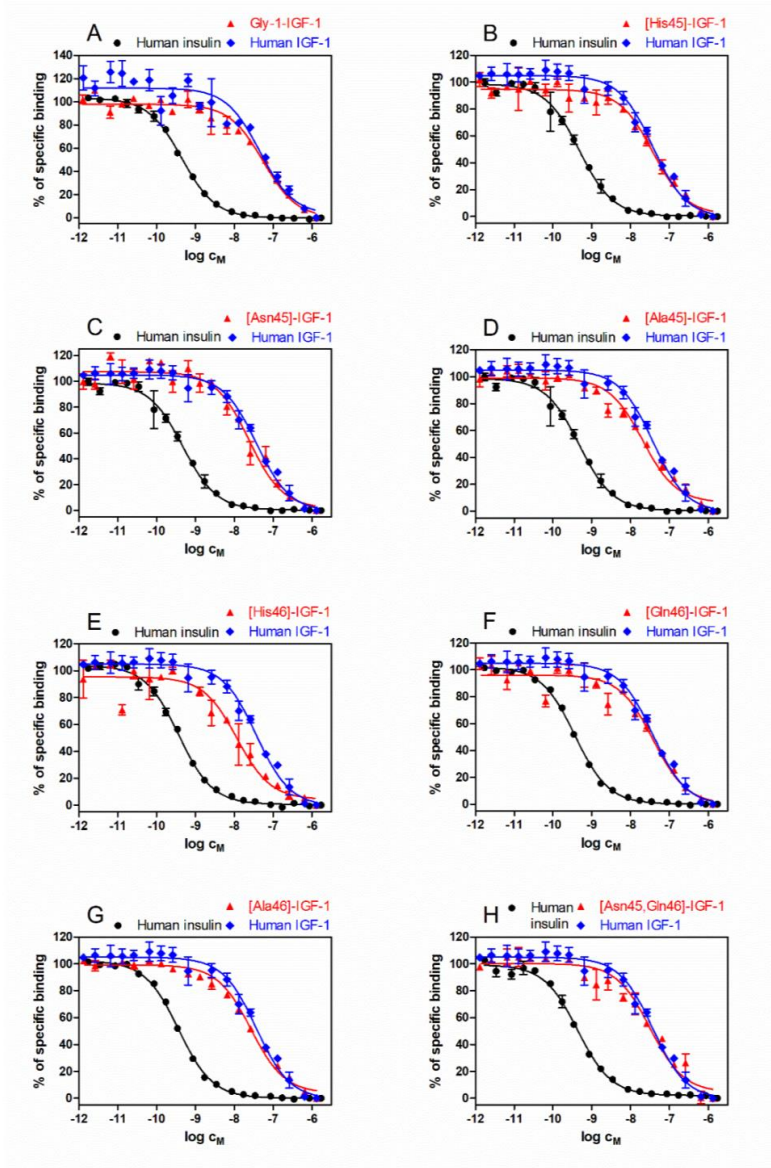


Figure S2. Inhibition of binding of human $[^{125}\text{I}]$ monoiodotyrosyl-insulin to human IR-A in human IM-9 lymphocytes by Gly-1-IGF-1 (A), [His45]-IGF-1 (B), [Asn45]-IGF-1 (C), [Ala45]-IGF-1 (D), [His46]-IGF-1 (E), [Gln46]-IGF-1 (F), [Ala46]-IGF-1 (G), [Asn45,Gln46]-IGF-1 (H) and by human insulin and human IGF-1 (in all panels). All IGF-1 analogs have Gly at the position -1. The representative binding curves are shown.

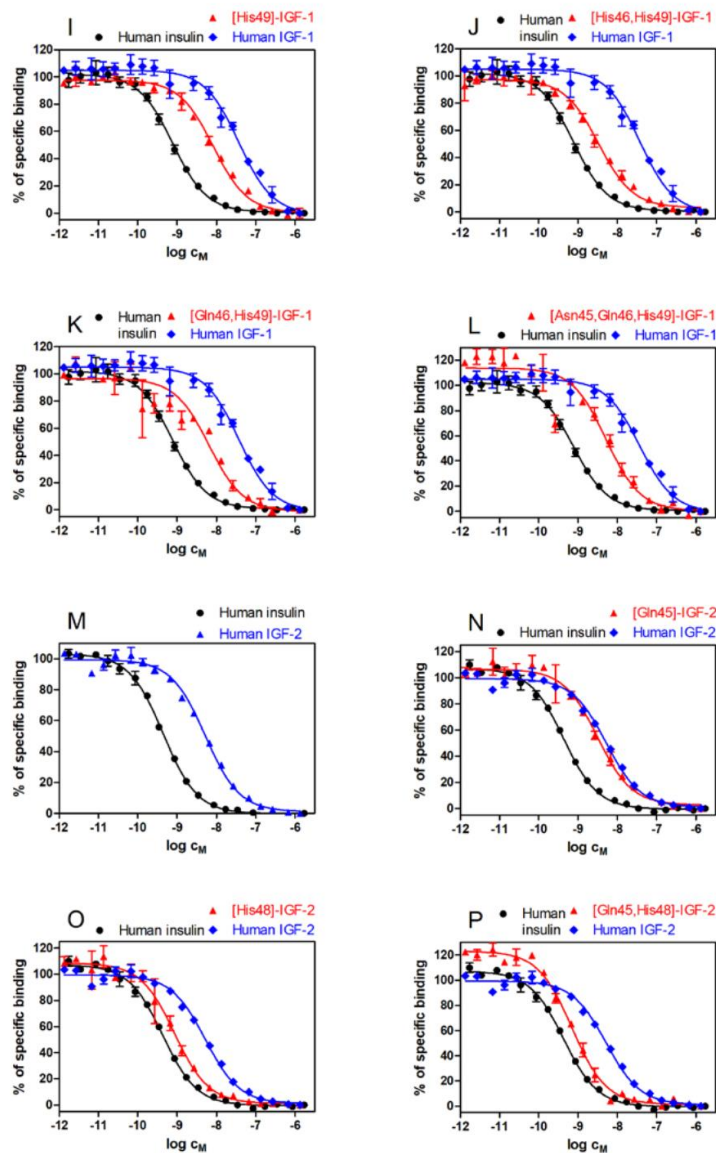


Figure S2 (continued). Inhibition of binding of human $[^{125}\text{I}]$ moniodotyrosyl-insulin to human IR-A in human IM-9 lymphocytes by $[\text{His}49]\text{-IGF-1}$ (**I**), $[\text{His}46,\text{His}49]\text{-IGF-1}$ (**J**), $[\text{Gln}46,\text{His}49]\text{-IGF-1}$ (**K**), $[\text{Asn}45,\text{Gln}46,\text{His}49]\text{-IGF-1}$ (**L**), human IGF-2 (**M**), $[\text{Gln}45]\text{-IGF-2}$ (**N**), $[\text{His}48]\text{-IGF-2}$ (**O**), $[\text{Gln}45,\text{His}48]\text{-IGF-2}$ (**P**) and by human insulin (in all panels), human IGF-1 (in panels **I-L**) and human IGF-2 (in panels **M-P**). All IGF-1 analogs have Gly at the position -1. The representative binding curves are shown.

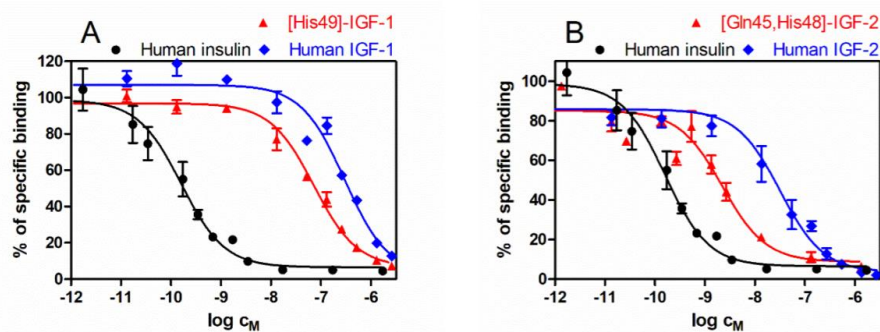


Figure S3. Inhibition of binding of human [¹²⁵I]moniodotyrosyl-insulin to human IR-B in mouse embryonic fibroblasts by human insulin, human IGF-1 and by [His49]-IGF-1 (in panel A), and by human insulin, human IGF-2 and by [Gln45,His48]-IGF-2 (in panel B). The representative binding curves are shown.



Figure S4. A representative example of immunoblot used for evaluation of abilities of analogs to induce the autophosphorylation of IGF-1R and IR-A at 10 nM concentration of the ligands.

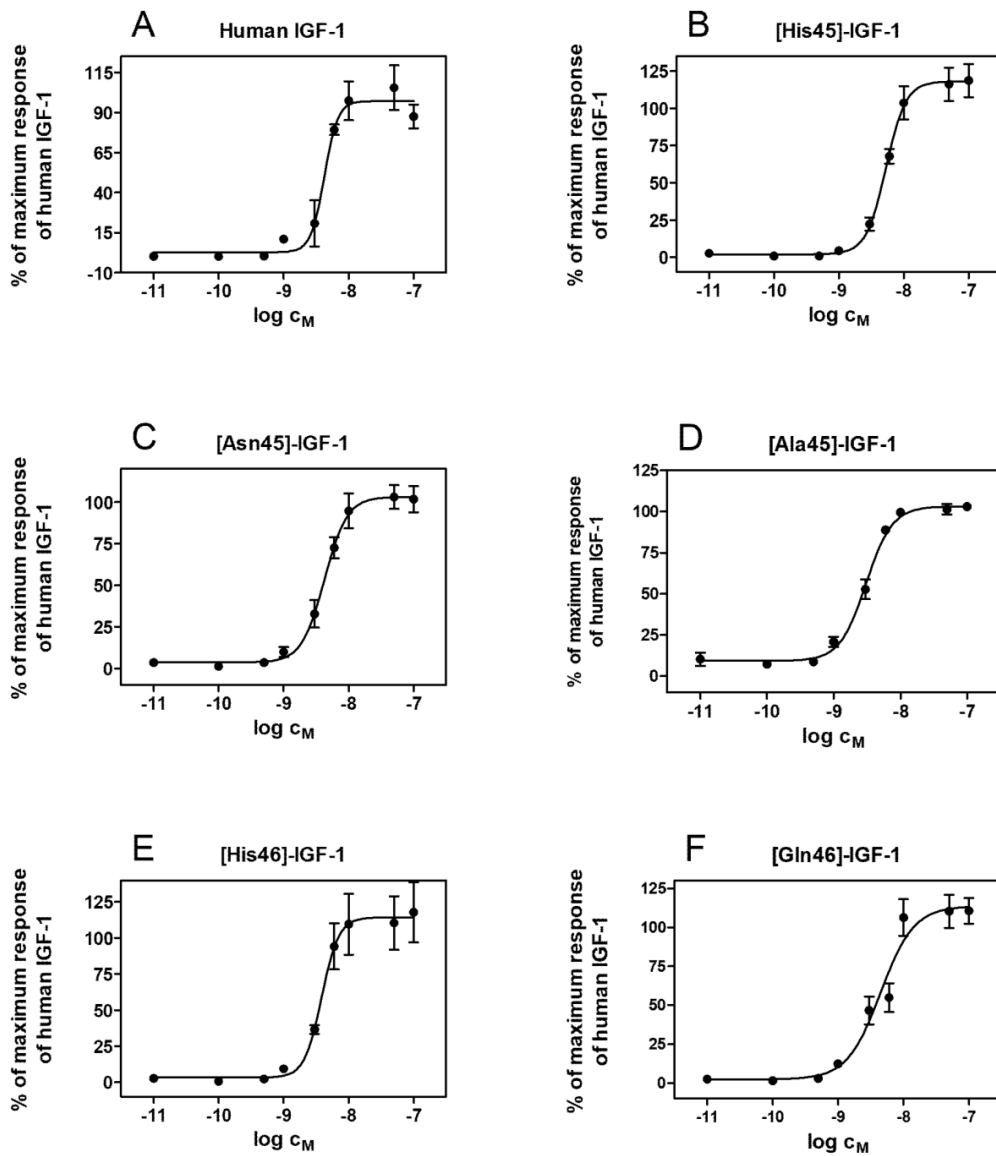


Figure S5. Representative dose response curves of the abilities of human IGF-1 and selected IGF-1 analogs to stimulate autophosphorylation of IGF-1R. Stimulation by human IGF-1 is shown in panel A, by [His45]-IGF-1 in panel B, by [Asn45]-IGF-1 in panel C, by [Ala45]-IGF-1 in panel D, by [His46]-IGF-1 in panel E and by [Gln45]-IGF-1 in panel F. The EC₅₀ values are shown in Table S2.

Table S2. Comparison of abilities of selected IGF-1 analogs to activate IGF-1R in a dose response manner (EC_{50} values) with their ability to activate IGF-1R at 10 nM concentration.

Analog	Ability to activate autophosphorylation of IGF-1R $EC_{50} \pm S.D.$ [nM] (n)	Relative ability ^a to activate autophosphorylation of IGF-1R [%]	Relative activation of IGF-1R at 10 nM [%] $\pm S.D.$ (n)
human IGF-1	3.7 ± 0.7 (3)	100 ± 19	100 ± 13 (4)
[His45]-IGF-1	5.4 ± 0.5 (3)	69 ± 6	71 ± 20 (4)
[Asn45]-IGF-1	4.7 ± 0.8 (3)	79 ± 13	88 ± 18 (4)
[Ala45]-IGF-1	3.1 ± 0.4 (3)	119 ± 15	80 ± 25 (4)
[His46]-IGF-1	4.4 ± 0.6 (4)	84 ± 11	71 ± 11 (4)
[Gln46]-IGF-1	4.4 ± 1.6 (3)	84 ± 30	69 ± 13 (4)

^a Relative ability to activate autophosphorylation of the receptor is defined as (EC_{50} of human IGF-1/ EC_{50} of analog) x 100.

References:

1. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension, *Nat. Protoc.* 2, 924-932.
2. Laible, M., and Boonrod, K. (2009) Homemade site directed mutagenesis of whole plasmids, *J. Vis. Exp.* 27, e1135, doi:10.3791/1135.

8 Discussion

8.1 Fully synthetic IGF-1 analogues with modifications at the positions 36 and 37.

8.1.1 *Development of the new methodology for the total chemical synthesis of IGF-1 analogues and preparation of the analogues.*

Total chemical synthesis of proteins, in contrast to the recombinant production of proteins, allows incorporation of non-coded amino acid and unusual structural motifs. On the other hand, the chemical synthesis of long (> 40-50 residues) peptides is a challenging task.

The chemical synthesis or semisynthesis is quite frequently used for the production of insulin analogues. Insulin consists of two peptide chains (A with 21 and B with 30 amino acids) and contains three disulphide bridges, two inter-chain between the A and B chains and one intra-chain in the A chain. A correct arrangement of disulphide bridges is the most difficult part of the synthesis [263]. A frequently used option is the use of orthogonal protection of cysteines (e.g. AcM-AcM, Trt-Trt and Npys-Npys) and the step-wise formation of respective disulphide pairs [264-269]. This is a robust but laborious protocol demanding purification of intermediates, which can result in relatively low yields.

An alternative simple and straightforward strategy benefits from the natural ability of insulin chains to adopt a native conformation and native combination of disulphide bridges. Similar process occurs in vivo in proinsulin. In our first variant of the chemical synthesis of insulin, the insulin chains A and B were protected with S-sulfonate groups [270-272]. This protection is stable at both basic and acidic conditions and therefore allows purification of the chains by HPLC. However, and remarkably, S-sulfonate group can be cleaved by a nearly stoichiometric amount of reducing agents (e.g. DTT). Next, a correct arrangement of disulphide bridges is done by slow oxidation of cysteines and a biomimetic folding. This method was used by Eli Lilly and Company in 80ies and 90ies for a production of tones of insulin. Another improvement of this oxidation can be achieved by folding of a prohormone form of insulin (i.e. with the A-chain N terminus and the B-chain C terminus connected by a peptide sequence or another A and B chain connective tag), followed by enzymatic removal of the connecting peptide, to give mature insulin [273]. Unfortunately, some specific mutations can damage the natural ability for correct biomimetic folding when using this method [271].

IGF-1, IGF-2 and proinsulin are single chain proteins, which are similar to insulin in primary and secondary structure, but they are significantly longer (70, 67 and 86 residues, respectively) than both insulin chains (21 and 30 residues). This number of residues leads to markedly lower efficiency of the solid-phase synthesis, as it was shown in reports on the synthesis of IGF-1 analogues [274-276]. Consequently, the total chemical synthesis is performed more frequently with insulin or other smaller hormones from the insulin protein family than on IGF-1, IGF-2 or proinsulin.

There are two elegant studies, which utilized a native chemical ligation of shorter protein fragments to overcome the solid-phase synthesis of the whole chains for preparation of proinsulin [277] or IGF-1 and its analogues [278].

In this thesis, we aimed to develop a new methodology for the total chemical synthesis of IGF-1 analogues based on the solid-phase synthesis of two shorter fragments with β -azido-alanine or propargyl glycine at the C- or N-terminus, respectively, their subsequent ligation by a Cu^{I} -catalyzed cycloaddition and a biomimetic recombination of disulphide bridges. As a result, IGF-1 analogues were connected by 1,2,3-triazole in the main chain, which is considered to have structural and electronic properties similar to those of the peptide bond [254]. To the best of our knowledge, our clicked IGF-1 analogues (relative molecular weights about 8,000 Da) are the longest disulphide-rich peptides with a triazole-link in the main chain prepared in a preparative scale.

The analogues prepared in this study are shown in Table 8.1. The 3D-fold of human IGF-1 (a crystal structure) is shown in Figure 8.1.

For the synthesis of fragments with β -azido-alanine we needed high amount of β -azido-alanine (a few grams). Fmoc β -azido L-alanine is commercially available, however its cost is high. Thus, we investigated the possibility of the in-house synthesis of Fmoc β -azido L-alanine and we succeeded to prepare grams of this compound in a reasonable time delay (the synthesis takes a week or two) and at user-friendly costs. Our Fmoc β -azido L-alanine was subsequently used for the synthesis of β -azido-alanine fragments.

Table 8.1. List of synthetic IGF-1 analogues prepared in this study.

Schematic structures	Analogue
$\text{H}_2\text{N-GPETLCGAELVDALQFVCGD RGFYFNKPTGYGSSSR-COOH}$ $\text{H}_2\text{N-RAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSA-COOH}$	1
$\text{H}_2\text{N-GPETLCGAELVDALQFVCGD RGFYFNKPTGYGSSS}$ $\text{HOOC-ASKAPKLPACYMELRRLDCSRFCCEDVIGTQPA}$	2
$\text{H}_2\text{N-GPETLCGAELVDALQFVCGD RGFYFNKPTGYGSSS}$ $\text{HOOC-ASKAPKLPACYNleELRRLDCSRFCCEDVIGTQPA}$	3
$\text{H}_2\text{N-GPETLCGAELVDALQFVCGD RGFYFNKPTGYGSSSR}$ $\text{HOOC-ASKAPKLPACYMELRRLDCSRFCCEDVIGTQPAR}$	4

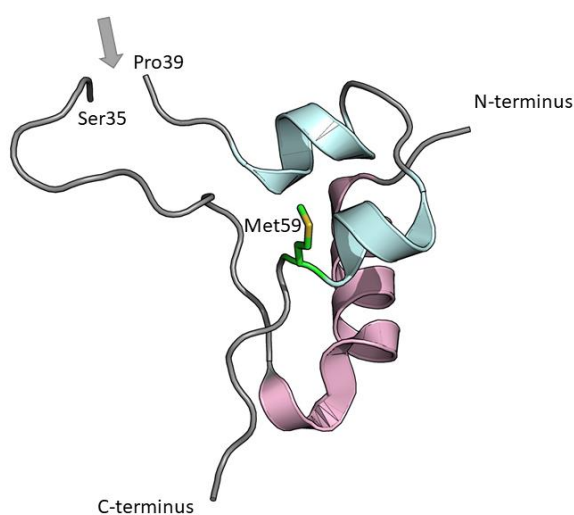


Figure 8.1. The 3D fold of human IGF-1. The interrupted chain (invisible residues 36-38), marked by a gray arrow, is the part of IGF-1 of our interest and was studied in this work. Electron density of this part of IGF-1 was missing in 1GZR.PDB crystal structure [13]. Side

chain of Met59 is shown in green, IGF-1 residues 39-66 are in light blue and residues 3-35 are in pink [279].

Solid-phase synthesis (performed by a Fmoc strategy) of one of the peptide chains for analogue **2**, was complicated by appearance of an uneasily separable unwanted-product, which was identified as methionine sulfoxide at the position 59. Therefore, we decided to prepare analogue **3**, where we substituted Met for Norleucine (Nle). Despite better purity of the crude product of the solid-phase synthesis, the synthetic yield was not improved, which led us back to native Met at position 59 in analogue **4**. Synthesis of the respective fragment for analogue **4** was accompanied by only a minor appearance of the oxidized product and it is not excluded, that the observed oxidation in the case of analogue **2** was just accidental due to minor errors during the manipulation with the crude peptide.

However, we observed rather important formation of an aspartimide derivative (8.6 %) during the synthesis of another IGF-1 peptide fragment having in the sequence the motif C¹⁸GDRG²². Appearance of this type of aspartimide is rather rare but not unusual [280]. This issue was overcome by the use of Fmoc-Asp(OEpe)-OH instead of Fmoc-Asp(OtBu)-OH, which suppressed the formation of the aspartimide.

Preparation of two chain IGF-1 (analogue **1**) was done by the synthesis of two peptides protected as *S*-sulphonates followed by their biomimetic combination by air oxidation (after the reduction to SH form). In Jiraček's laboratory this procedure is used routinely as a straightforward methodology for the synthesis of insulin analogues [271, 272]. The yields of this synthetic step for human insulin are typically about 10 % [271]. However, in the case of two chain IGF-1 the yield was only about 0.6 %. This indicates the importance of the single chain structure of IGF-1 for a proper folding of the hormone.

Using the same *S*-sulphonate protecting strategy for cysteines, we attempted to obtain one chain IGF-1 analogue **2** by Cu^I-catalyzed cycloaddition ligation of two shorter fragments with C-terminal β-azido alanine (chain 1-36) and N-terminal propargyl glycine (fragment 37-70). However, all our attempts for “clicking” these fragments by Cu^I-catalyzed click reaction failed and we have never isolated any products. First of all we used as a catalyst CuSO₄ reduced by ascorbic acid in numerous solvents, since we had a positive experience with these conditions [281]. Further, was used as catalyst CuBr with TBTA or THPTA, as accelerating ligands [258], in various solvents, including 5 M urea or 6 M guanidine hydrochloride (as a chaotropic agent in water) or 0.5 M tetrabutylammonium bromide (as a chaotropic agent in organic solvents). However, all these reactions failed. We concluded that

S-sulphonate groups are incompatible with copper-catalyzed azide-alkyne cycloaddition, probably because of the ability of S-sulphonate to chelate copper, which could lead to undesirable Cu-mediated folding of IGF-1 chains and thus to steric blocking of azido and alkyne groups.

Hence, in further experiments we chose an alternative acetamidomethyl protection of cysteines since AcM-protection should not interfere with the clicking step and offers several different chemical ways for the deprotection [282]. The cycloaddition of the AcM-protected chains was effective only with a high excess of CuSO₄/ascorbate as a catalyst in water-tBuOH. Addition of click accelerators [258], which should work as Cu^I stabilizing chelating agents, or reactions in organic solvents did not lead to desired products. Interestingly, reactions were not effective if scaled up above 2-4 mg per chain. It seems that formation of complexes with Cu^I causes steric problems during cycloaddition reactions with large peptide precursors. The failure of the reaction in organic solvents could be due to a necessity of an effective protein hydration, which is better in a mixture of water and tBuOH than in organic solvents only. The reactions performed with precursors without arginines gave better yields than with precursors containing arginines. This could be due to some steric shielding of azido and alkyne moieties by bulky arginine side chains.

The moderate yields of the ligation reactions (7-18 %) document the difficulty of the click reactions in case of large peptides or proteins as precursors compared to click reactions with small peptides or small molecules, which are usually quick and almost quantitative [254].

The deprotection of acetamidomethyl groups from cysteines and the folding of analogues (forming of S-S bridges), were crucial for the success of the synthesis. Deprotection by an excess of iodine in acetic acid, frequently reported in the literature [269], was not successful. In our case, the majority of AcM groups remained uncleaved even after longer reaction time (more than 4 hours). Therefore, we used a procedure of AcM group deprotection with silver trifluoromethanesulphonate in TFA [283] followed by application of gaseous hydrogen sulfide to remove Ag⁺ from cysteine salts. Finally, the precursors were allowed to slowly air-oxidize and fold spontaneously. This three-step methodology provided yields about 6-8 % for analogues **2** and **3**, which is comparable to our approximately 10 % yield of a single step folding of human insulin from its A and B chains [271] and much better than for two chain IGF-1 analogue **1** (0.6 %). These results document more effective folding of single chain IGF-1 precursors, than two chain IGF-1 analogues. The folding of analogue

4 was markedly less effective (1.7 %) than folding of analogues **2** and **3**. The reason could be the presence of arginines at the positions 36 and 39 of the analogue causes extension a C-domain, we actually insert of an extra triazole-containing connecting moiety, which has a negative effect on the folding of the IGF-1.

8.1.2 Receptor binding and activation properties of fully synthetic IGF-1 analogues.

Mutation of the pair of arginines Arg36-Arg37 for Ala36-Ala37 severely impaired binding of the mutant IGF-1 to IGF-1R [19, 20, 23, 39] and to IGFBP-1 [21]. Furthermore, two studies indicated that the analogue with two glutamates instead of Arg36-Arg37 can have features like a partial antagonist, i.e. still effective binding to IGF-1R but impaired receptor signaling [261] and tumorigenesis [284].

Moreover, a missing electron density of residues 36-38 (Arg-Arg-Ala) in the crystal structure of human IGF-1 [13] led to the theory if the pair of basic arginines 36 and 37 could be a cleavage site for proteolytic enzymes *in vivo*, and whether the cleavage of Arg36-Arg37 peptide bond can disrupt binding to IGF-1R or whether the resulting “two chain” IGF-1 is still able to effectively bind and activate IGF-1R.

The binding affinity of the two chain IGF-1 analogue **1** for IGF-1R was severely impaired (3.5 %) compared to native IGF-1. This result does not support the theory that the cleavage of Arg36-Arg37 peptide bond (if it happens *in vivo*) is another step of IGF-1 prohormone processing. However, most likely, it is a degradation step leading to a much less active protein.

The binding affinity of analogue **2** for IGF-1R is 21 % of native IGF-1, which is less pronounced drop than the effect of the Arg36Ala/Arg37Ala mutation (about 5 %) [20]. Interestingly, it seems that the triazole ring, which is three atoms longer than the peptide bond in native IGF-1 (see Figure 8.2.), and neighboring ethyl carboxyamido (at 36) and ethyl amino (at 37) moieties, are somehow compensated the absence of arginine residues.

The analogue **3** differs from analogue **2** only in Nle residue at the position **59** instead of Met, and had affinity reduced to 5 % of the affinity of native IGF-1. This observation indicates the importance of sulfur atom in Met59 for IGF-1R binding affinity of IGF-1. In agreement with that is the 17-fold decrease in binding of Phe59 IGF-1 [28].

The highest IGF-1R binding affinity in this series has analogue **4**, about 25 % of native IGF-1. This analogue contains C-loop arginines at positions 36 and 39 instead at

positions 36 and 37. Their C α atoms are separated by eight extra backbone atoms plus triazole ring (see Figure 8.2.). The moderate increase of binding affinity of this analogue (compared to analogue **2**) indicates a relative tolerance of the receptor for prolongation of this part of the C-domain. In addition, it is possible that minimum one of the arginine residues (36 or 39) occupies similar position on the receptor as in the case of native hormone.

These results could open new possibilities for the use of different side chains and motifs at positions adjacent to the triazole ring, i.e. 36 and 37 in **2** and 37 and 38 in **4**. However, such advancement depends on the development of available functionalized azides and alkynes suitable also for the solid-phase peptide synthesis.

All analogues **1-4** have reduced binding affinities for IR-A compared to native IGF-1. However, the decrease is less pronounced than for IGF-1R (21-43 % of native IGF-1). This could indicate that the C-domain of IGF-1 is more important for binding to IGF-1R than for IR-A.

All analogues activated autophosphorylation of IGF-1R and IR-A proportionally to their binding affinities for these receptors without any signs of potential antagonism. Therefore, previously observed [261] partial antagonism of Glu36-Glu37 IGF-1 can be probably dependent on negative charges of glutamic acid side chains.

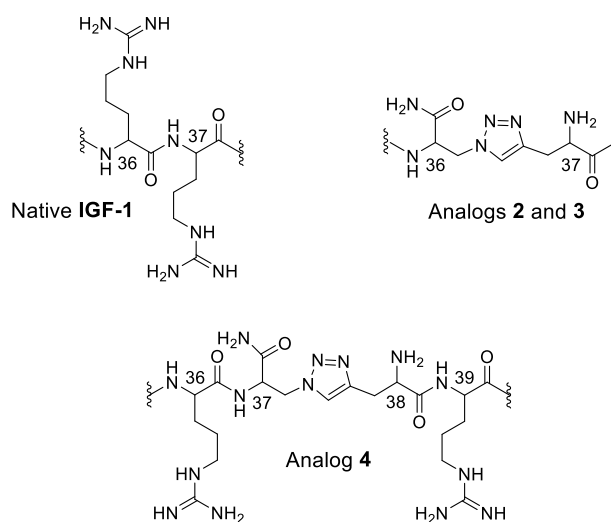


Figure 8.2. The comparison of the “connecting” motifs at positions 36-37 or 37-38 of analogues 2-4 with native IGF-1. C α atoms are marked by residue numbers [279].

8.2 Development of a new methodology for the recombinant production of IGF-1, IGF-2 and their analogues in *E. coli*.

Our aim was to develop a new efficient, straightforward and cost-effective protocol for the recombinant production of IGF-1, IGF-2 and their analogues in *E. coli* [49, 259].

Novo Nordisk A/S routinely uses recombinant production of proinsulin in *Saccharomyces cerevisiae*, which is followed by enzymatic processing to form mature insulin [285]. However, because respective expression vectors are not easily available, we decided to use *E. coli* for the production and folding of single chain IGFs employing a methodology routinely used in IOCB for the production of other proteins [286]. We started with IGF-2 because it is much less studied hormone than IGF-1 and because we supposed that the methodology developed for IGF-2 will work also for IGF-1.

IGF-2 was expressed as a fusion with an N-terminal, and cleavable by TEV protease, His6-tagged GB1 protein (immunoglobulin binding domain B1 of Streptococcal Protein-G) [286], which increases solubility and allows as fast and easy purification on Ni-column.

The crucial point of the purification protocol was the cleavage of the fusion partner from IGF-2 precursor and subsequent formation of mature IGF-2 while avoiding misfolded variant of the hormone. We used TEV protease, which needs reducing conditions to be properly active. However, reductive conditions strongly damage proper folding of IGF-2. After optimizations steps, we found convenient conditions for both proper cleavage by TEV protease and proper IGF-2 formation in a mixture of oxidized and reduced glutathione (specifically 1.5 mM GSH, 0.15 mM GSSG). Interestingly, moderately reducing environment triggered a correct forming of disulphide bonds of the peptide.

Since this technique provided relatively high yields (0.5-1.8 mg L⁻¹ of culture) of IGF-2 we decided to customize this methodology for the production of IGF-1. IGF-2 can be produced without any additional amino acids at the N-terminus using Glu-Asn-Leu-Tyr-Phe-Gln↓Ala TEV protease cleavage sequence. The cleavage at (↓) results in the native IGF-2 hormone with N-terminal alanine. Unfortunately, this strategy is not successful for the IGF-1 due to the Pro residue at the position 2 in IGF-1 (Figure 5.2.), which does not allow TEV protease to cut the cleavage sequence. Therefore, for the production of IGF-1 we inserted an extra Gly residue at the -1 position of the hormone, which enabled the cleavage of the precursor. Consequently, all our IGF-1 analogues had an additional Gly at -1 position.

Further biological characterization of Gly-1-IGF-1 showed that this extra N-terminal glycine has no effect on the binding affinity to IGF-1R and IR receptors.

8.3 Synthesis and characterization of IGF-1 analogues with mutations at the positions 45, 46 and 49.

Since upregulation of IGF-1 action plays a role in development of cancer, a lot of efforts have been focused on the development of effective anti-IGF-1R-directed therapies. The majority of these compounds were tyrosine-kinase (TK) inhibitors and anti-receptor antibodies [287]. Unfortunately, the results of clinical trials were not satisfactory [247], because of the toxicity and an increasing overlap and takeover of IGF-1R signalling pathways by the IR. The new anticancer therapies specially target only IGF-1R receptor, i.e. IGF-based selective antagonist of the IGF-1R, should represent a promising alternative strategy to combat IGF-1R-related malignancies. To date, no IGF-like peptide antagonists of the IGF-1R have been identified. However peptides with antagonistic properties were discovered by a phage-display technique [288].

Whittaker et al. demonstrated that a combination of GluA4His and ThrA8His mutations of insulin results in insulin analogues with native IR binding affinity, but with a poor signalling transduction. They also proposed, based on mutational and IR photo crosslinking experiments, that surfaces involving insulin GluA4 and IR Asp707 could play a crucial role in the mechanism of receptor activation [262]. Insulin acidic GluA4 is preserved by its equivalent Asp45 in the IGF-1 and insulin-neutral GlnA5 is replaced by Glu46 in corresponding sites of IGF-1 (see Figure 8.3.). Furthermore, Asp 707 of IR α -CT is replaced by a neutral Asn694 in the IGF-1R α -CT segment. These correlations and trends between positions A4/A5 in insulin and IGF-1 equivalent 45/46 sites led us to study impact of the mutations of these residues on receptor affinities and autophosphorylation in order to develop IGF-1R-specific antagonists with potential anti-cancer clinical applications. We systematically modified IGF-1 Asp45 site with Asn or Ala and Glu46 site with Gln or Ala. Phe49 was modified with His because equivalent mutation in insulin dramatically increased its binding affinity for IR-A (see Figure 8.3.). Afterwards we characterized IGF-1R and IR-A binding and autophosphorylation properties of these analogues.

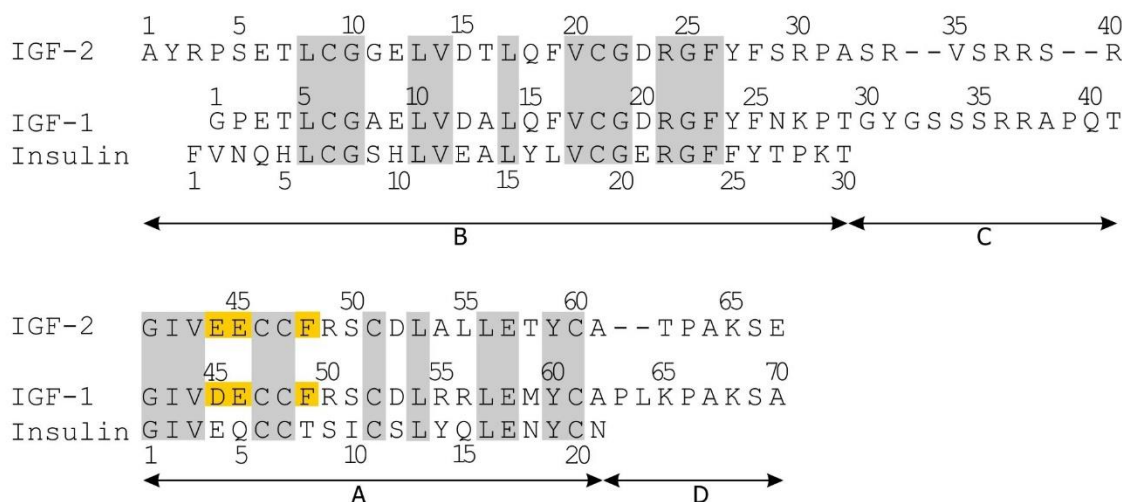


Figure 8.3. Comparison of the primary sequences of human IGF-1, IGF-2 and insulin with highlighted mutated sites. The homologous regions are shown with a gray background and the mutated residues are highlighted with a yellow background. Reproduced from Ref. [49].

We did not discover any clear and major antagonism or receptor overstimulation. All analogues activated IR-A and IGF-1R without major disproportions to their binding affinities. Therefore, the previously described significant disproportionate binding and activation of insulin analogues mutated at A4/A8 by Whittaker et al. [262] is probably specific and limited only for insulin-IR interactions.

Mutations at the position 45 led to reduce binding affinities (29-68 %) to IGF-1R, while mutations at the position 46 led to the same, or slightly higher, binding affinities (92-133 %) to IGF-1R. These analogues have binding potencies for IR-A similar to the native IGF-1.

Much more surprising was the effect of mutation of Phe49 in IGF-1 for His. This mutation enhanced binding to IR-A about 5-fold, which makes the respective analogues the strongest IGF-1 binders of the IR-A receptor isoform thus far reported. Similar, about 3-fold enhanced binding affinity was determined for IR-B as well. Thus, the mutation is specific for IR receptors only and not for IGF-1R, which may indicate that the respective binding sites (probably parts of receptors' binding sites 2) are not equivalent. Similar effects were observed for equivalent HisA8 in insulin [262, 289-291] and His48 in IGF-2 (this study). In addition, the same mutation in IGF-2 (Phe48His) slightly enhances binding affinity of the analogue for IGF-1R, which is not the case for the mutation Phe49His in IGF-1. This may indicate that the interactions of IGF-1 and IGF-2 with IGF-1R are different. The residues at

position 49 of IGF-1 and 48 of IGF-2 have an important role in IGF-1R interaction and do not interact with the receptor in the same fashion.

Some trends observed in the analogues described here can be corroborated by the phylogeny of insulin-related hormones. Although there is no mammalian insulin with His at site A8, this amino acid is frequently present there in fish, frog and bird insulins [292]. Also, HisA8 was proposed as being responsible for about a 5-fold higher binding affinity of chicken insulin for IR in human lymphocytes [293]. Herring et al. [294] proposed that it is possible that a “lower affinity” ThrA8 site in mammals emerged from the evolutionary optimization of the insulin-receptor kinetics, which requires a reduction of insulin binding affinity in mammals that is dictated by a different route of insulin delivery. In more ancient vertebrates like birds, insulin is secreted differently, by kidney, which leads to rapid metabolic changes [294]. However, in mammals, the pancreas-to-liver portal vein-mediated pathway is a primary direction of insulin. In humans, the liver is the main glycemic-response organ with >90 % “metabolic” IR-B isoform, and where insulin inhibits gluconeogenesis and glycogenolysis, before reaching the peripheral tissues. It may be postulated that to assure an optimum liver:body distribution of insulin its IR-binding affinity needs to be reduced in mammals, in order to achieve a more systemic glycemic response [294].

9 Conclusions

We developed a new methodology for the total chemical synthesis of IGF-1 analogues. The methodology is based on ligation of two precursor chains prepared by solid-phase peptide synthesis, by subsequent Cu^I-catalyzed cycloaddition and finally by a biomimetic formation of three disulphide bridges. It is an alternative synthetic strategy to previously established methods using native ligation and orthogonal protection of cysteine residues. The “clicked” IGF-1 analogues reported here are unique examples of the use of triazole-forming click chemistry for the preparative production of large peptides/small proteins (8,000 Da) with a complicated disulphide pattern. Connection of two IGF-1 fragments by a triazole-containing moieties was relatively well tolerated in binding to IGF-1R and IR-A, and can thus represent a suitable synthetic platform for the design of new analogues. In addition, we found that the disconnection of IGF-1 main chain at 36-37 peptide bond leads to a deactivation of the hormone, which indicates that the respective cleavage is not a part of the maturation of the prohormone *in vivo*.

We developed a straightforward protocol for the production of the recombinant IGF-1 with one extra N-terminal Gly(-1) and prepared eleven IGF-1 analogues with mutations at positions 45, 46 and 49, which were subsequently tested for receptor binding affinities and abilities activation to activate their cognate receptors. We did not detect any clear and major antagonism or receptor overstimulation. However, we found that mutation Phe49His in IGF-1 dramatically enhances binding affinities for IR-A and IR-B, which makes these mutant the strongest IR-A/IR-B binders reported so far. Moreover, this mutation does not change binding affinity of analogues for IGF-1R indicating that respective sites (probably sites 2) in IRs and IGF-1R receptors are not equivalent.

This work brought new valuable information about structure-activity studies of IGF-1 and IGF-1R and opened new directions for the further rational engineering of the hormones of the insulin-IGF system.

10 References

1. Denley, A., et al., *Molecular interactions of the IGF system*. Cytokine & Growth Factor Reviews, 2005. **16**(4): p. 421-439.
2. Clemmons, D.R., *Modifying IGF1 activity: an approach to treat endocrine disorders, atherosclerosis and cancer*. Nat Rev Drug Discov, 2007. **6**(10): p. 821-33.
3. Adams, T.E., et al., *Structure and function of the type I insulin-like growth factor receptor*. Cell Mol Life Sci, 2000. **57**(7): p. 1050-93.
4. Sciacca, L., R. Le Moli, and R. Vigneri, *Insulin analogs and cancer*. Front Endocrinol (Lausanne), 2012. **3**: p. 21.
5. Yaseen, M.A., K.C. Pedley, and S.L. Howell, *Regulation of insulin secretion from islets of Langerhans rendered permeable by electric discharge*. Biochemical Journal, 1982. **206**(1): p. 81-87.
6. Rinderknecht, E. and R.E. Humbel, *The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin*. J Biol Chem, 1978. **253**(8): p. 2769-76.
7. Sato, A., et al., *Three-dimensional structure of human insulin-like growth factor-I (IGF-I) determined by 1H-NMR and distance geometry*. Int J Pept Protein Res, 1993. **41**(5): p. 433-40.
8. Schaffer, M.L., et al., *Complex with a phage display-derived peptide provides insight into the function of insulin-like growth factor I*. Biochemistry, 2003. **42**(31): p. 9324-34.
9. Cooke, R.M., T.S. Harvey, and I.D. Campbell, *Solution structure of human insulin-like growth factor I: a nuclear magnetic resonance and restrained molecular dynamics study*. Biochemistry, 1991. **30**(22): p. 5484-91.
10. Laajoki, L.G., et al., *Solution structure and backbone dynamics of long-[Arg(3)]insulin-like growth factor-I*. J Biol Chem, 2000. **275**(14): p. 10009-15.
11. Vajdos, F.F., et al., *Crystal structure of human insulin-like growth factor-I: detergent binding inhibits binding protein interactions*. Biochemistry, 2001. **40**(37): p. 11022-9.
12. Zeslawski, W., et al., *The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5*. Embo j, 2001. **20**(14): p. 3638-44.
13. Brzozowski, A.M., et al., *Structural origins of the functional divergence of human insulin-like growth factor-I and insulin*. Biochemistry, 2002. **41**(30): p. 9389-97.
14. Siwanowicz, I., et al., *Structural basis for the regulation of insulin-like growth factors by IGF binding proteins*. Structure, 2005. **13**(1): p. 155-67.
15. Torres, A.M., et al., *Solution structure of human insulin-like growth factor II. Relationship to receptor and binding protein interactions*. J Mol Biol, 1995. **248**(2): p. 385-401.
16. Smith, G.D., W.A. Pangborn, and R.H. Blessing, *The structure of T6 human insulin at 1.0 Å resolution*. Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 3): p. 474-82.
17. Jiráček, J. and L. Žáková, *Structural Perspectives of Insulin Receptor Isoform-Selective Insulin Analogs*. Frontiers in Endocrinology, 2017. **8**(167).
18. Pandini, G., et al., *Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved*. J Biol Chem, 2002. **277**(42): p. 39684-95.

19. Denley, A., et al., *Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR*. Mol Endocrinol, 2004. **18**(10): p. 2502-12.
20. Zhang, W., et al., *Positively charged side chains in the insulin-like growth factor-I C- and D-regions determine receptor binding specificity*. J Biol Chem, 1994. **269**(14): p. 10609-13.
21. Jansson, M., et al., *The insulin-like growth factor (IGF)binding protein 1 binding epitope on IGF-I probed by heteronuclear NMR spectroscopy and mutational analysis*. J Biol Chem, 1998. **273**(38): p. 24701-7.
22. Bayne, M.L., et al., *The roles of tyrosines 24, 31, and 60 in the high affinity binding of insulin-like growth factor-I to the type 1 insulin-like growth factor receptor*. J Biol Chem, 1990. **265**(26): p. 15648-52.
23. Bayne, M.L., et al., *The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor*. J Biol Chem, 1989. **264**(19): p. 11004-8.
24. Kristensen, C., et al., *A single-chain insulin-like growth factor I/insulin hybrid binds with high affinity to the insulin receptor*. Biochem J, 1995. **305** (Pt 3): p. 981-6.
25. Cara, J.F., et al., *An insulin-like growth factor I/insulin hybrid exhibiting high potency for interaction with the type I insulin-like growth factor and insulin receptors of placental plasma membranes*. J Biol Chem, 1990. **265**(29): p. 17820-5.
26. Bagley, C.J., et al., *A key functional role for the insulin-like growth factor I N-terminal pentapeptide*. Biochem J, 1989. **259**(3): p. 665-71.
27. Bayne, M.L., et al., *Structural analogs of human insulin-like growth factor I with reduced affinity for serum binding proteins and the type 2 insulin-like growth factor receptor*. J Biol Chem, 1988. **263**(13): p. 6233-9.
28. Shooter, G.K., et al., *Insulin-like growth factor (IGF)-I A- and B-domain analogues with altered type 1 IGF and insulin receptor binding specificities*. J Mol Endocrinol, 1996. **17**(3): p. 237-46.
29. Dubaquié, Y., et al., *Binding protein-3-selective insulin-like growth factor I variants: engineering, biodistributions, and clearance*. Endocrinology, 2001. **142**(1): p. 165-73.
30. King, R., et al., *Production and characterization of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu3, following their expression in Escherichia coli as fusion proteins*. J Mol Endocrinol, 1992. **8**(1): p. 29-41.
31. De Meyts, P. and J. Whittaker, *Structural biology of insulin and IGF1 receptors: implications for drug design*. Nat Rev Drug Discov, 2002. **1**(10): p. 769-83.
32. Magee, B.A., et al., *Insulin-like growth factor I and its binding proteins: a study of the binding interface using B-domain analogues*. Biochemistry, 1999. **38**(48): p. 15863-70.
33. Schwartz, G.P., G.T. Burke, and P.G. Katsoyannis, *A superactive insulin: [B10-aspartic acid]insulin(human)*. Proc Natl Acad Sci U S A, 1987. **84**(18): p. 6408-11.
34. Denley, A., et al., *Structural and functional characteristics of the Val44Met insulin-like growth factor I missense mutation: correlation with effects on growth and development*. Mol Endocrinol, 2005. **19**(3): p. 711-21.
35. Gauguin, L., et al., *Alanine scanning of a putative receptor binding surface of insulin-like growth factor-I*. J Biol Chem, 2008. **283**(30): p. 20821-9.

36. Walenkamp, M.J., et al., *Homozygous and heterozygous expression of a novel insulin-like growth factor-I mutation*. J Clin Endocrinol Metab, 2005. **90**(5): p. 2855-64.
37. Forbe, B.E., et al., *Contribution of residues A54 and L55 of the human insulin-like growth factor-II (IGF-II) A domain to Type 2 IGF receptor binding specificity*. Growth Factors, 2001. **19**(3): p. 163-73.
38. Cascieri, M.A., et al., *Structural analogs of human insulin-like growth factor (IGF) I with altered affinity for type 2 IGF receptors*. J Biol Chem, 1989. **264**(4): p. 2199-202.
39. Gill, R., et al., *Engineering the C-region of human insulin-like growth factor-1: implications for receptor binding*. Protein Eng, 1996. **9**(11): p. 1011-9.
40. Gauguin, L., et al., *Structural basis for the lower affinity of the insulin-like growth factors for the insulin receptor*. J Biol Chem, 2008. **283**(5): p. 2604-13.
41. Francis, G.L., et al., *Insulin-like growth factor (IGF)-II binding to IGF-binding proteins and IGF receptors is modified by deletion of the N-terminal hexapeptide or substitution of arginine for glutamate-6 in IGF-II*. Biochem J, 1993. **293** (Pt 3): p. 713-9.
42. Milner, S.J., et al., *Mutations in the B-domain of insulin-like growth factor-I influence the oxidative folding to yield products with modified biological properties*. Biochemical Journal, 1995. **308**(Pt 3): p. 865-871.
43. Forbes, B.E., et al., *Characteristics of binding of insulin-like growth factor (IGF)-I and IGF-II analogues to the type 1 IGF receptor determined by BIAcore analysis*. Eur J Biochem, 2002. **269**(3): p. 961-8.
44. Hodgson, D.R., F.E. May, and B.R. Westley, *Involvement of phenylalanine 23 in the binding of IGF-I to the insulin and type I IGF receptor*. Regul Pept, 1996. **66**(3): p. 191-6.
45. Hodgson, D.R., F.E. May, and B.R. Westley, *Mutations at positions 11 and 60 of insulin-like growth factor I reveal differences between its interactions with the type I insulin-like-growth-factor receptor and the insulin receptor*. Eur J Biochem, 1995. **233**(1): p. 299-309.
46. Jansson, M., M. Uhlen, and B. Nilsson, *Structural changes in insulin-like growth factor (IGF) I mutant proteins affecting binding kinetic rates to IGF binding protein 1 and IGF-I receptor*. Biochemistry, 1997. **36**(14): p. 4108-17.
47. Cascieri, M.A., et al., *Mutants of human insulin-like growth factor I with reduced affinity for the type 1 insulin-like growth factor receptor*. Biochemistry, 1988. **27**(9): p. 3229-33.
48. Francis, G.L., et al., *Novel recombinant fusion protein analogues of insulin-like growth factor (IGF)-I indicate the relative importance of IGF-binding protein and receptor binding for enhanced biological potency*. J Mol Endocrinol, 1992. **8**(3): p. 213-23.
49. Macháčková, K., et al., *Converting Insulin-like Growth Factors 1 and 2 into High-Affinity Ligands for Insulin Receptor Isoform A by the Introduction of an Evolutionarily Divergent Mutation*. Biochemistry, 2018. **57**(16): p. 2373-2382.
50. Clemmons, D.R., et al., *Competition for binding to insulin-like growth factor (IGF) binding protein-2, 3, 4, and 5 by the IGFs and IGF analogs*. Endocrinology, 1992. **131**(2): p. 890-5.
51. Dubaquié, Y. and H.B. Lowman, *Total alanine-scanning mutagenesis of insulin-like growth factor I (IGF-I) identifies differential binding epitopes for IGFBP-1 and IGFBP-3*. Biochemistry, 1999. **38**(20): p. 6386-96.

52. Bach, L.A., et al., *Binding of mutants of human insulin-like growth factor II to insulin-like growth factor binding proteins 1-6*. J Biol Chem, 1993. **268**(13): p. 9246-54.
53. Forbes, B., et al., *Classification of the insulin-like growth factor binding proteins into three distinct categories according to their binding specificities*. Biochem Biophys Res Commun, 1988. **157**(1): p. 196-202.
54. Clemmons, D.R., et al., *Discrete alterations of the insulin-like growth factor I molecule which alter its affinity for insulin-like growth factor-binding proteins result in changes in bioactivity*. J Biol Chem, 1990. **265**(21): p. 12210-6.
55. Yakar, S., et al., *Normal growth and development in the absence of hepatic insulin-like growth factor I*. Proc Natl Acad Sci U S A, 1999. **96**(13): p. 7324-9.
56. Yee, D., *Targeting insulin-like growth factor pathways*. Br J Cancer, 2007. **96 Suppl**: p. R7-10.
57. Isgaard, J., et al., *Regulation of insulin-like growth factor messenger ribonucleic acid in rat growth plate by growth hormone*. Endocrinology, 1988. **122**(4): p. 1515-20.
58. Clemmons, D.R., *Value of Insulin-like Growth Factor System Markers in the Assessment of Growth Hormone Status*. Endocrinology and Metabolism Clinics of North America, 2007. **36**(1): p. 109-129.
59. Powell-Braxton, L., et al., *IGF-I is required for normal embryonic growth in mice*. Genes Dev, 1993. **7**(12b): p. 2609-17.
60. Ye, P., et al., *Astrocyte-specific overexpression of insulin-like growth factor-I promotes brain overgrowth and glial fibrillary acidic protein expression*. J Neurosci Res, 2004. **78**(4): p. 472-84.
61. Liu, J.P., et al., *Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type I IGF receptor (Igf1r)*. Cell, 1993. **75**(1): p. 59-72.
62. Lissett, C.A., H. Gleeson, and S.M. Shalet, *The insulin-like growth factor I generation test in adults*. Horm Res, 2004. **62 Suppl 1**: p. 44-9.
63. Hartman, M.L., et al., *A low dose euglycemic infusion of recombinant human insulin-like growth factor I rapidly suppresses fasting-enhanced pulsatile growth hormone secretion in humans*. J Clin Invest, 1993. **91**(6): p. 2453-62.
64. Clemmons, D.R. and J.J. Van Wyk, *Factors controlling blood concentration of somatomedin C*. Clin Endocrinol Metab, 1984. **13**(1): p. 113-43.
65. Cercek, B., et al., *Induction of insulin-like growth factor I messenger RNA in rat aorta after balloon denudation*. Circ Res, 1990. **66**(6): p. 1755-60.
66. Edwall, D., et al., *Induction of insulin-like growth factor I messenger ribonucleic acid during regeneration of rat skeletal muscle*. Endocrinology, 1989. **124**(2): p. 820-5.
67. Chu, C.R., et al., *Recovery of articular cartilage metabolism following thermal stress is facilitated by IGF-I and JNK inhibitor*. Am J Sports Med, 2004. **32**(1): p. 191-6.
68. O'Donnell, S.L., et al., *IGF-I and microglia/macrophage proliferation in the ischemic mouse brain*. Glia, 2002. **39**(1): p. 85-97.
69. Swantek, J.L. and R. Baserga, *Prolonged activation of ERK2 by epidermal growth factor and other growth factors requires a functional insulin-like growth factor I receptor*. Endocrinology, 1999. **140**(7): p. 3163-9.
70. Clemmons, D.R. and L.E. Underwood, *Role of insulin-like growth factors and growth hormone in reversing catabolic states*. Horm Res, 1992. **38 Suppl 2**: p. 37-40.

71. Fryburg, D.A., *Insulin-like growth factor I exerts growth hormone- and insulin-like actions on human muscle protein metabolism*. Am J Physiol, 1994. **267**(2 Pt 1): p. E331-6.
72. Bark, T.H., et al., *Increased protein synthesis after acute IGF-I or insulin infusion is localized to muscle in mice*. Am J Physiol, 1998. **275**(1 Pt 1): p. E118-23.
73. Párrizas, M., et al., *Abundant Insulin-like Growth Factor-1 (IGF-1) Receptor Binding in Fish Skeletal Muscle*. General and Comparative Endocrinology, 1995. **98**(1): p. 16-25.
74. Kuzuya, H., et al., *Trial of insulinlike growth factor I therapy for patients with extreme insulin resistance syndromes*. Diabetes, 1993. **42**(5): p. 696-705.
75. Fernandez, A.M., et al., *Functional inactivation of the IGF-I and insulin receptors in skeletal muscle causes type 2 diabetes*. Genes Dev, 2001. **15**(15): p. 1926-34.
76. Yakar, S., et al., *Inhibition of growth hormone action improves insulin sensitivity in liver IGF-1-deficient mice*. J Clin Invest, 2004. **113**(1): p. 96-105.
77. Kappeler, L., et al., *Brain IGF-1 receptors control mammalian growth and lifespan through a neuroendocrine mechanism*. PLoS Biol, 2008. **6**(10): p. e254.
78. Holzenberger, M., L. Kappeler, and C.D.M. Filho, *IGF-1 signaling and aging*. Experimental Gerontology, 2004. **39**(11): p. 1761-1764.
79. Walenkamp, M.J., M. Losekoot, and J.M. Wit, *Molecular IGF-1 and IGF-1 receptor defects: from genetics to clinical management*. Endocr Dev, 2013. **24**: p. 128-37.
80. Hankinson, S.E., et al., *Circulating concentrations of insulin-like growth factor-I and risk of breast cancer*. Lancet, 1998. **351**(9113): p. 1393-6.
81. Ranke, M.B., *Insulin-like growth factor-I treatment of growth disorders, diabetes mellitus and insulin resistance*. Trends Endocrinol Metab, 2005. **16**(4): p. 190-7.
82. Backeljauw, P.F. and L.E. Underwood, *Therapy for 6.5-7.5 years with recombinant insulin-like growth factor I in children with growth hormone insensitivity syndrome: a clinical research center study*. J Clin Endocrinol Metab, 2001. **86**(4): p. 1504-10.
83. Chernausk, S.D., et al., *Long-term treatment with recombinant insulin-like growth factor (IGF)-I in children with severe IGF-I deficiency due to growth hormone insensitivity*. J Clin Endocrinol Metab, 2007. **92**(3): p. 902-10.
84. Lang, C.H., L. Hong-Brown, and R.A. Frost, *Cytokine inhibition of JAK-STAT signaling: a new mechanism of growth hormone resistance*. Pediatr Nephrol, 2005. **20**(3): p. 306-12.
85. Mauras, N. and B. Beaufriere, *Recombinant human insulin-like growth factor-I enhances whole body protein anabolism and significantly diminishes the protein catabolic effects of prednisone in humans without a diabetogenic effect*. J Clin Endocrinol Metab, 1995. **80**(3): p. 869-74.
86. Debroy, M.A., et al., *Anabolic effects of insulin-like growth factor in combination with insulin-like growth factor binding protein-3 in severely burned adults*. J Trauma, 1999. **47**(5): p. 904-10; discussion 910-1.
87. Davis, M.E., et al., *Local myocardial insulin-like growth factor I (IGF-1) delivery with biotinylated peptide nanofibers improves cell therapy for myocardial infarction*. Proc Natl Acad Sci U S A, 2006. **103**(21): p. 8155-60.
88. Grinspoon, S., et al., *Effects of short-term recombinant human insulin-like growth factor I administration on bone turnover in osteopenic women with anorexia nervosa*. J Clin Endocrinol Metab, 1996. **81**(11): p. 3864-70.
89. Boonen, S., et al., *Musculoskeletal effects of the recombinant human IGF-I/IGF binding protein-3 complex in osteoporotic patients with proximal femoral fracture:*

- a double-blind, placebo-controlled pilot study.* J Clin Endocrinol Metab, 2002. **87**(4): p. 1593-9.
90. Khorsandi, M., et al., *Effects of hypophysectomy on vascular insulin-like growth factor-I gene expression after balloon denudation in rats.* Atherosclerosis, 1992. **93**(1-2): p. 115-22.
 91. Zhu, B., et al., *Targeted Overexpression of IGF-I in Smooth Muscle Cells of Transgenic Mice Enhances Neointimal Formation through Increased Proliferation and Cell Migration after Intraarterial Injury.* Endocrinology, 2001. **142**(8): p. 3598-3606.
 92. Thrailkill, K.M., et al., *Cotherapy with recombinant human insulin-like growth factor I and insulin improves glycemic control in type 1 diabetes. RhIGF-I in IDDM Study Group.* Diabetes Care, 1999. **22**(4): p. 585-92.
 93. Clemmons, D.R., et al., *Rh/IGF-I/rhIGFBP-3 administration to patients with type 2 diabetes mellitus reduces insulin requirements while also lowering fasting glucose.* Growth Hormone & IGF Research, 2005. **15**(4): p. 265-274.
 94. Moses, A.C., et al., *Recombinant human insulin-like growth factor I increases insulin sensitivity and improves glycemic control in type II diabetes.* Diabetes, 1996. **45**(1): p. 91-100.
 95. Holzenberger, M., et al., *IGF-I receptor regulates lifespan and resistance to oxidative stress in mice.* Nature, 2003. **421**(6919): p. 182-7.
 96. Kenyon, C., *A conserved regulatory system for aging.* Cell, 2001. **105**(2): p. 165-8.
 97. Ogg, S., et al., *The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans.* Nature, 1997. **389**(6654): p. 994-9.
 98. Clancy, D.J., et al., *Extension of life-span by loss of CHICO, a Drosophila insulin receptor substrate protein.* Science, 2001. **292**(5514): p. 104-6.
 99. Rincon, M., E. Rudin, and N. Barzilai, *The insulin/IGF-I signaling in mammals and its relevance to human longevity.* Experimental Gerontology, 2005. **40**(11): p. 873-877.
 100. Talbot, K., et al., *Demonstrated brain insulin resistance in Alzheimer's disease patients is associated with IGF-I resistance, IRS-1 dysregulation, and cognitive decline.* The Journal of Clinical Investigation, 2012. **122**(4): p. 1316-1338.
 101. Humbert, S., et al., *The IGF-I/Akt Pathway Is Neuroprotective in Huntington's Disease and Involves Huntingtin Phosphorylation by Akt.* Developmental Cell, 2002. **2**(6): p. 831-837.
 102. Firth, S.M. and R.C. Baxter, *Cellular actions of the insulin-like growth factor binding proteins.* Endocr Rev, 2002. **23**(6): p. 824-54.
 103. Storr, H.L., et al., *Heterogeneity of the growth phenotype and birth size in acid-labile subunit (ALS) deficiency.* J Endocrinol Invest, 2015. **38**(4): p. 407-12.
 104. Boisclair, Y.R., et al., *The acid-labile subunit (ALS) of the 150 kDa IGF-binding protein complex: an important but forgotten component of the circulating IGF system.* J Endocrinol, 2001. **170**(1): p. 63-70.
 105. Twigg, S.M. and R.C. Baxter, *Insulin-like growth factor (IGF)-binding protein 5 forms an alternative ternary complex with IGFs and the acid-labile subunit.* J Biol Chem, 1998. **273**(11): p. 6074-9.
 106. Jones, J.I. and D.R. Clemmons, *Insulin-like growth factors and their binding proteins: biological actions.* Endocr Rev, 1995. **16**(1): p. 3-34.
 107. Boes, M., et al., *Insulin-like growth factor binding protein (IGFBP)4 accounts for the connective tissue distribution of endothelial cell IGFBPs perfused through the isolated heart.* Endocrinology, 1992. **131**(1): p. 327-30.

108. Lewitt, M.S., et al., *Complex formation by human insulin-like growth factor-binding protein-3 and human acid-labile subunit in growth hormone-deficient rats.* Endocrinology, 1994. **134**(6): p. 2404-9.
109. Wheatcroft, S.B. and M.T. Kearney, *IGF-dependent and IGF-independent actions of IGF-binding protein-1 and -2: implications for metabolic homeostasis.* Trends Endocrinol Metab, 2009. **20**(4): p. 153-62.
110. Bach, L.A., S.J. Headey, and R.S. Norton, *IGF-binding proteins--the pieces are falling into place.* Trends Endocrinol Metab, 2005. **16**(5): p. 228-34.
111. Neumann, G.M. and L.A. Bach, *The N-terminal disulfide linkages of human insulin-like growth factor-binding protein-6 (hIGFBP-6) and hIGFBP-1 are different as determined by mass spectrometry.* J Biol Chem, 1999. **274**(21): p. 14587-94.
112. Forbes, B.E., P. McCarthy, and R.S. Norton, *Insulin-like growth factor binding proteins: a structural perspective.* Front Endocrinol (Lausanne), 2012. **3**: p. 38.
113. Kalus, W., et al., *Structure of the IGF-binding domain of the insulin-like growth factor-binding protein-5 (IGFBP-5): implications for IGF and IGF-I receptor interactions.* The EMBO Journal, 1998. **17**(22): p. 6558-6572.
114. Forbes, B.E., et al., *Localization of an insulin-like growth factor (IGF) binding site of bovine IGF binding protein-2 using disulfide mapping and deletion mutation analysis of the C-terminal domain.* J Biol Chem, 1998. **273**(8): p. 4647-52.
115. Galanis, M., et al., *Ligand-binding characteristics of recombinant amino- and carboxyl-terminal fragments of human insulin-like growth factor-binding protein-3.* J Endocrinol, 2001. **169**(1): p. 123-33.
116. Carrick, F.E., B.E. Forbes, and J.C. Wallace, *BIAcore analysis of bovine insulin-like growth factor (IGF)-binding protein-2 identifies major IGF binding site determinants in both the amino- and carboxyl-terminal domains.* J Biol Chem, 2001. **276**(29): p. 27120-8.
117. Firth, S.M., U. Ganeshprasad, and R.C. Baxter, *Structural determinants of ligand and cell surface binding of insulin-like growth factor-binding protein-3.* J Biol Chem, 1998. **273**(5): p. 2631-8.
118. Marinaro, J.A., et al., *O-glycosylation of insulin-like growth factor (IGF) binding protein-6 maintains high IGF-II binding affinity by decreasing binding to glycosaminoglycans and susceptibility to proteolysis.* Eur J Biochem, 2000. **267**(17): p. 5378-86.
119. Marinaro, J.A., D.J. Casley, and L.A. Bach, *O-glycosylation delays the clearance of human IGF-binding protein-6 from the circulation.* Eur J Endocrinol, 2000. **142**(5): p. 512-6.
120. Firth, S.M. and R.C. Baxter, *Characterisation of recombinant glycosylation variants of insulin-like growth factor binding protein-3.* J Endocrinol, 1999. **160**(3): p. 379-87.
121. Dolcini, L., et al., *Identification of the amniotic fluid insulin-like growth factor binding protein-1 phosphorylation sites and propensity to proteolysis of the isoforms.* Febs j, 2009. **276**(20): p. 6033-46.
122. Gibson, J.M., et al., *Regulation of IGF bioavailability in pregnancy.* Mol Hum Reprod, 2001. **7**(1): p. 79-87.
123. Coverley, J.A., J.L. Martin, and R.C. Baxter, *The effect of phosphorylation by casein kinase 2 on the activity of insulin-like growth factor-binding protein-3.* Endocrinology, 2000. **141**(2): p. 564-70.

124. Schedlich, L.J., et al., *Phosphorylation of insulin-like growth factor binding protein-3 by deoxyribonucleic acid-dependent protein kinase reduces ligand binding and enhances nuclear accumulation*. *Endocrinology*, 2003. **144**(5): p. 1984-93.
125. Rajah, R., et al., *Insulin-like growth factor binding protein (IGFBP) proteases: functional regulators of cell growth*. *Prog Growth Factor Res*, 1995. **6**(2-4): p. 273-84.
126. Lalou, C., C. Lassarre, and M. Binoux, *Isolation and characterization of proteolytic fragments of insulin-like growth factor-binding protein-3*. *Horm Res*, 1996. **45**(3-5): p. 156-9.
127. Giudice, L.C., et al., *Insulin-like growth factor binding proteins in maternal serum throughout gestation and in the puerperium: effects of a pregnancy-associated serum protease activity*. *J Clin Endocrinol Metab*, 1990. **71**(4): p. 806-16.
128. Hossenlopp, P., et al., *Evidence of enzymatic degradation of insulin-like growth factor-binding proteins in the 150K complex during pregnancy*. *J Clin Endocrinol Metab*, 1990. **71**(4): p. 797-805.
129. Cwyfan Hughes, S.C., et al., *The induction of specific proteases for insulin-like growth factor-binding proteins following major heart surgery*. *J Endocrinol*, 1992. **135**(1): p. 135-45.
130. Davies, S.C., et al., *The induction of a specific protease for insulin-like growth factor binding protein-3 in the circulation during severe illness*. *J Endocrinol*, 1991. **130**(3): p. 469-73.
131. Lemmon, M.A. and J. Schlessinger, *Cell Signaling by Receptor Tyrosine Kinases*. *Cell*, 2010. **141**(7): p. 1117-1134.
132. Rubin, J.B., M.A. Shia, and P.F. Pilch, *Stimulation of tyrosine-specific phosphorylation in vitro by insulin-like growth factor I*. *Nature*, 1983. **305**(5933): p. 438-40.
133. Ullrich, A. and J. Schlessinger, *Signal transduction by receptors with tyrosine kinase activity*. *Cell*, 1990. **61**(2): p. 203-12.
134. Ullrich, A., et al., *Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity*. *Embo j*, 1986. **5**(10): p. 2503-12.
135. Abbott, A.M., et al., *Insulin-like growth factor I receptor gene structure*. *J Biol Chem*, 1992. **267**(15): p. 10759-63.
136. Jacobs, S., F.C. Kull, Jr., and P. Cuatrecasas, *Monensin blocks the maturation of receptors for insulin and somatomedin C: identification of receptor precursors*. *Proc Natl Acad Sci U S A*, 1983. **80**(5): p. 1228-31.
137. Bhaumick, B., R.M. Bala, and M.D. Hollenberg, *Somatomedin receptor of human placenta: solubilization, photolabeling, partial purification, and comparison with insulin receptor*. *Proc Natl Acad Sci U S A*, 1981. **78**(7): p. 4279-83.
138. Chernausk, S.D., S. Jacobs, and J.J. Van Wyk, *Structural similarities between human receptors for somatomedin C and insulin: analysis by affinity labeling*. *Biochemistry*, 1981. **20**(26): p. 7345-50.
139. Lawrence, M.C., N.M. McKern, and C.W. Ward, *Insulin receptor structure and its implications for the IGF-1 receptor*. *Curr Opin Struct Biol*, 2007. **17**(6): p. 699-705.
140. Olson, T.S., M.J. Bamberger, and M.D. Lane, *Post-translational changes in tertiary and quaternary structure of the insulin proreceptor. Correlation with acquisition of function*. *J Biol Chem*, 1988. **263**(15): p. 7342-51.

141. Ward, C.W., P.A. Hoyne, and R.H. Flegg, *Insulin and epidermal growth factor receptors contain the cysteine repeat motif found in the tumor necrosis factor receptor*. *Proteins*, 1995. **22**(2): p. 141-53.
142. Ward, C.W., *Members of the insulin receptor family contain three fibronectin type III domains*. *Growth Factors*, 1999. **16**(4): p. 315-22.
143. Marino-Buslje, C., et al., *A third fibronectin type III domain in the extracellular region of the insulin receptor family*. *FEBS Lett*, 1998. **441**(2): p. 331-6.
144. Mulhern, T.D., G.W. Booker, and L. Cosgrove, *A third fibronectin-type-III domain in the insulin-family receptors*. *Trends Biochem Sci*, 1998. **23**(12): p. 465-6.
145. Garrett, T.P., et al., *Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor*. *Nature*, 1998. **394**(6691): p. 395-9.
146. Favelyukis, S., et al., *Structure and autoregulation of the insulin-like growth factor I receptor kinase*. *Nat Struct Biol*, 2001. **8**(12): p. 1058-63.
147. Pautsch, A., et al., *Crystal Structure of Bisphosphorylated IGF-1 Receptor Kinase*. *Structure*, 2001. **9**(10): p. 955-965.
148. Anastassiadis, T., et al., *A highly selective dual insulin receptor (IR)/insulin-like growth factor I receptor (IGF-1R) inhibitor derived from an extracellular signal-regulated kinase (ERK) inhibitor*. *J Biol Chem*, 2013. **288**(39): p. 28068-77.
149. Wittman, M.D., et al., *Discovery of a 2,4-disubstituted pyrrolo[1,2-f][1,2,4]triazine inhibitor (BMS-754807) of insulin-like growth factor receptor (IGF-1R) kinase in clinical development*. *J Med Chem*, 2009. **52**(23): p. 7360-3.
150. Menting, J.G., et al., *How insulin engages its primary binding site on the insulin receptor*. *Nature*, 2013. **493**(7431): p. 241-5.
151. Ward, C.W. and M.C. Lawrence, *Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor*. *Bioessays*, 2009. **31**(4): p. 422-34.
152. Smith, B.J., et al., *Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists*. *Proc Natl Acad Sci U S A*, 2010. **107**(15): p. 6771-6.
153. Kiselyov, V.V., et al., *Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation*. *Mol Syst Biol*, 2009. **5**: p. 243.
154. Schaffer, L., *A model for insulin binding to the insulin receptor*. *Eur J Biochem*, 1994. **221**(3): p. 1127-32.
155. Menting, J.G., et al., *Protective hinge in insulin opens to enable its receptor engagement*. *Proc Natl Acad Sci U S A*, 2014. **111**(33): p. E3395-404.
156. Žáková, L., et al., *Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex*. *Acta Crystallographica Section D: Biological Crystallography*, 2014. **70**(Pt 10): p. 2765-2774.
157. Zakova, L., et al., *Structural integrity of the B24 site in human insulin is important for hormone functionality*. *J Biol Chem*, 2013. **288**(15): p. 10230-40.
158. Menting, J.G., et al., *Structural Congruency of Ligand Binding to the Insulin and Insulin/Type 1 Insulin-like Growth Factor Hybrid Receptors*. *Structure*, 2015. **23**(7): p. 1271-82.
159. Whittaker, L., et al., *High-affinity insulin binding: insulin interacts with two receptor ligand binding sites*. *Biochemistry*, 2008. **47**(48): p. 12900-9.
160. Scapin, G., et al., *Structure of the insulin receptor-insulin complex by single-particle cryo-EM analysis*. *Nature*, 2018. **556**(7699): p. 122-125.

161. Xu, Y., et al., *How ligand binds to the type 1 insulin-like growth factor receptor*. Nature Communications, 2018. **9**: p. 821.
162. Gronborg, M., et al., *Structure-function relationship of the insulin-like growth factor-I receptor tyrosine kinase*. J Biol Chem, 1993. **268**(31): p. 23435-40.
163. Kato, H., et al., *Essential role of tyrosine residues 1131, 1135, and 1136 of the insulin-like growth factor-I (IGF-I) receptor in IGF-I action*. Mol Endocrinol, 1994. **8**(1): p. 40-50.
164. Li, S., et al., *Mitogenicity and transforming activity of the insulin-like growth factor-I receptor with mutations in the tyrosine kinase domain*. J Biol Chem, 1994. **269**(51): p. 32558-64.
165. Heldin, C.H., *Dimerization of cell surface receptors in signal transduction*. Cell, 1995. **80**(2): p. 213-23.
166. Girnita, L., et al., *Something old, something new and something borrowed: emerging paradigm of insulin-like growth factor type 1 receptor (IGF-1R) signaling regulation*. Cell Mol Life Sci, 2014. **71**(13): p. 2403-27.
167. Jones, N., et al., *Analysis of a Shc family adaptor protein, ShcD/Shc4, that associates with muscle-specific kinase*. Mol Cell Biol, 2007. **27**(13): p. 4759-73.
168. Craparo, A., T.J. O'Neill, and T.A. Gustafson, *Non-SH2 domains within insulin receptor substrate-1 and SHC mediate their phosphotyrosine-dependent interaction with the NPEY motif of the insulin-like growth factor I receptor*. J Biol Chem, 1995. **270**(26): p. 15639-43.
169. Dey, B.R., et al., *Evidence for the direct interaction of the insulin-like growth factor I receptor with IRS-1, Shc, and Grb10*. Mol Endocrinol, 1996. **10**(6): p. 631-41.
170. Sasaoka, T., et al., *Comparison of the insulin and insulin-like growth factor I mitogenic intracellular signaling pathways*. Endocrinology, 1996. **137**(10): p. 4427-34.
171. He, W., et al., *Interaction of insulin receptor substrate-2 (IRS-2) with the insulin and insulin-like growth factor I receptors. Evidence for two distinct phosphotyrosine-dependent interaction domains within IRS-2*. J Biol Chem, 1996. **271**(20): p. 11641-5.
172. Fukushima, T., et al., *Insulin receptor substrates form high-molecular-mass complexes that modulate their availability to insulin/insulin-like growth factor-I receptor tyrosine kinases*. Biochem Biophys Res Commun, 2011. **404**(3): p. 767-73.
173. Yoshihara, H., et al., *Insulin/insulin-like growth factor (IGF) stimulation abrogates an association between a deubiquitinating enzyme USP7 and insulin receptor substrates (IRSs) followed by proteasomal degradation of IRSs*. Biochem Biophys Res Commun, 2012. **423**(1): p. 122-7.
174. Takahashi, Y., et al., *Roles of insulin receptor substrate-1 and Shc on insulin-like growth factor I receptor signaling in early passages of cultured human fibroblasts*. Endocrinology, 1997. **138**(2): p. 741-50.
175. Chardin, P., et al., *Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2*. Science, 1993. **260**(5112): p. 1338-43.
176. Ward, C.W., et al., *Systematic mapping of potential binding sites for Shc and Grb2 SH2 domains on insulin receptor substrate-1 and the receptors for insulin, epidermal growth factor, platelet-derived growth factor, and fibroblast growth factor*. J Biol Chem, 1996. **271**(10): p. 5603-9.
177. Alessi, D.R., et al., *Identification of the sites in MAP kinase kinase-1 phosphorylated by p74raf-1*. Embo j, 1994. **13**(7): p. 1610-9.

178. Plotnikov, A., et al., *The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation*. Biochim Biophys Acta, 2011. **1813**(9): p. 1619-33.
179. McCubrey, J.A., et al., *Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance*. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2007. **1773**(8): p. 1263-1284.
180. Myers, M.G., Jr., et al., *IRS-1 is a common element in insulin and insulin-like growth factor-I signaling to the phosphatidylinositol 3'-kinase*. Endocrinology, 1993. **132**(4): p. 1421-30.
181. Fukushima, T., et al., *Phosphatidylinositol 3-kinase (PI3K) activity bound to insulin-like growth factor-I (IGF-I) receptor, which is continuously sustained by IGF-I stimulation, is required for IGF-I-induced cell proliferation*. J Biol Chem, 2012. **287**(35): p. 29713-21.
182. Balendran, A., et al., *Evidence that 3-phosphoinositide-dependent protein kinase-1 mediates phosphorylation of p70 S6 kinase in vivo at Thr-412 as well as Thr-252*. J Biol Chem, 1999. **274**(52): p. 37400-6.
183. Alessi, D.R., et al., *Mechanism of activation of protein kinase B by insulin and IGF-I*. The EMBO Journal, 1996. **15**(23): p. 6541-6551.
184. del Peso, L., et al., *Interleukin-3-induced phosphorylation of BAD through the protein kinase Akt*. Science, 1997. **278**(5338): p. 687-9.
185. Datta, S.R., et al., *Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery*. Cell, 1997. **91**(2): p. 231-41.
186. Cardone, M.H., et al., *Regulation of cell death protease caspase-9 by phosphorylation*. Science, 1998. **282**(5392): p. 1318-21.
187. Sehat, B., et al., *Identification of c-Cbl as a new ligase for insulin-like growth factor-I receptor with distinct roles from Mdm2 in receptor ubiquitination and endocytosis*. Cancer Res, 2008. **68**(14): p. 5669-77.
188. Carelli, S., et al., *Degradation of insulin-like growth factor-I receptor occurs via ubiquitin-proteasome pathway in human lung cancer cells*. J Cell Physiol, 2006. **208**(2): p. 354-62.
189. Segretin, M.E., et al., *Insulin-like growth factor-1 receptor regulation in activated human T lymphocytes*. Horm Res, 2003. **59**(6): p. 276-80.
190. Girnita, L., A. Girnita, and O. Larsson, *Mdm2-dependent ubiquitination and degradation of the insulin-like growth factor 1 receptor*. Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8247-52.
191. Vecchione, A., et al., *The Grb10/Nedd4 complex regulates ligand-induced ubiquitination and stability of the insulin-like growth factor I receptor*. Mol Cell Biol, 2003. **23**(9): p. 3363-72.
192. Girnita, L., et al., *{beta}-Arrestin is crucial for ubiquitination and down-regulation of the insulin-like growth factor-1 receptor by acting as adaptor for the MDM2 E3 ligase*. J Biol Chem, 2005. **280**(26): p. 24412-9.
193. Sarfstein, R., et al., *Insulin-like growth factor-I receptor (IGF-IR) translocates to nucleus and autoregulates IGF-IR gene expression in breast cancer cells*. J Biol Chem, 2012. **287**(4): p. 2766-76.
194. Aleksic, T., et al., *Type 1 insulin-like growth factor receptor translocates to the nucleus of human tumor cells*. Cancer Res, 2010. **70**(16): p. 6412-9.
195. Sehat, B., et al., *SUMOylation mediates the nuclear translocation and signaling of the IGF-1 receptor*. Sci Signal, 2010. **3**(108): p. ra10.

196. Deng, H., et al., *Over-accumulation of nuclear IGF-1 receptor in tumor cells requires elevated expression of the receptor and the SUMO-conjugating enzyme Ubc9*. *Biochem Biophys Res Commun*, 2011. **404**(2): p. 667-71.
197. Sarfstein, R. and H. Werner, *Minireview: nuclear insulin and insulin-like growth factor-1 receptors: a novel paradigm in signal transduction*. *Endocrinology*, 2013. **154**(5): p. 1672-9.
198. Warsito, D., et al., *Nuclear IGF1R is a transcriptional co-activator of LEF1/TCF*. *EMBO Rep*, 2012. **13**(3): p. 244-50.
199. Schlessinger, J. and M.A. Lemmon, *Nuclear signaling by receptor tyrosine kinases: the first robin of spring*. *Cell*, 2006. **127**(1): p. 45-8.
200. Benyoucef, S., et al., *Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation*. *Biochemical Journal*, 2007. **403**(Pt 3): p. 603-613.
201. Slaaby, R., et al., *Hybrid receptors formed by insulin receptor (IR) and insulin-like growth factor I receptor (IGF-IR) have low insulin and high IGF-1 affinity irrespective of the IR splice variant*. *J Biol Chem*, 2006. **281**(36): p. 25869-74.
202. Bailyes, E.M., et al., *Insulin receptor/IGF-1 receptor hybrids are widely distributed in mammalian tissues: quantification of individual receptor species by selective immunoprecipitation and immunoblotting*. *Biochem J*, 1997. **327** (Pt 1): p. 209-15.
203. Soos, M.A., et al., *Receptors for insulin and insulin-like growth factor-I can form hybrid dimers. Characterisation of hybrid receptors in transfected cells*. *Biochemical Journal*, 1990. **270**(2): p. 383-390.
204. Seely, B.L., et al., *A functional assessment of insulin/insulin-like growth factor-I hybrid receptors*. *Endocrinology*, 1995. **136**(4): p. 1635-41.
205. Belfiore, A., et al., *Insulin/IGF-1 hybrid receptors play a major role in IGF-1 signaling in thyroid cancer*. *Biochimie*, 1999. **81**(4): p. 403-407.
206. Slaaby, R., *Specific insulin/IGF1 hybrid receptor activation assay reveals IGF1 as a more potent ligand than insulin*. *Sci Rep*, 2015. **5**: p. 7911.
207. Samani, A.A. and P. Brodt, *The receptor for the type I insulin-like growth factor and its ligands regulate multiple cellular functions that impact on metastasis*. *Surg Oncol Clin N Am*, 2001. **10**(2): p. 289-312, viii.
208. Baserga, R., M. Resnicoff, and M. Dews, *The IGF-1 receptor and cancer*. *Endocrine*, 1997. **7**(1): p. 99-102.
209. Chan, J.M., et al., *Plasma insulin-like growth factor-I and prostate cancer risk: a prospective study*. *Science*, 1998. **279**(5350): p. 563-6.
210. Yu, H., et al., *Plasma levels of insulin-like growth factor-I and lung cancer risk: a case-control analysis*. *J Natl Cancer Inst*, 1999. **91**(2): p. 151-6.
211. Petridou, E., et al., *Endometrial cancer and the IGF system: a case-control study in Greece*. *Oncology*, 2003. **64**(4): p. 341-5.
212. Pietrkowski, Z., et al., *Inhibition of growth of prostatic cancer cell lines by peptide analogues of insulin-like growth factor I*. *Cancer Res*, 1993. **53**(5): p. 1102-6.
213. Sepp-Lorenzino, L., et al., *Herbimycin A induces the 20 S proteasome- and ubiquitin-dependent degradation of receptor tyrosine kinases*. *J Biol Chem*, 1995. **270**(28): p. 16580-7.
214. Navab, R., et al., *Inhibition of endosomal insulin-like growth factor-I processing by cysteine proteinase inhibitors blocks receptor-mediated functions*. *J Biol Chem*, 2001. **276**(17): p. 13644-9.

215. Burtrum, D., et al., *A fully human monoclonal antibody to the insulin-like growth factor I receptor blocks ligand-dependent signaling and inhibits human tumor growth in vivo*. *Cancer Res*, 2003. **63**(24): p. 8912-21.
216. Beltran, P.J., et al., *Efficacy of ganitumab (AMG 479), alone and in combination with rapamycin, in Ewing's and osteogenic sarcoma models*. *J Pharmacol Exp Ther*, 2011. **337**(3): p. 644-54.
217. Zia, F., et al., *Monoclonal antibody alpha IR-3 inhibits non-small cell lung cancer growth in vitro and in vivo*. *J Cell Biochem Suppl*, 1996. **24**: p. 269-75.
218. Goetsch, L., et al., *A recombinant humanized anti-insulin-like growth factor receptor type I antibody (h7C10) enhances the antitumor activity of vinorelbine and anti-epidermal growth factor receptor therapy against human cancer xenografts*. *Int J Cancer*, 2005. **113**(2): p. 316-28.
219. Lu, D., et al., *A fully human recombinant IgG-like bispecific antibody to both the epidermal growth factor receptor and the insulin-like growth factor receptor for enhanced antitumor activity*. *J Biol Chem*, 2005. **280**(20): p. 19665-72.
220. Olmos, D., et al., *Targeting the Insulin-Like Growth Factor I Receptor in Ewing's Sarcoma: Reality and Expectations*. *Sarcoma*, 2011. **2011**: p. 402508.
221. Olmos, D., et al., *Safety, pharmacokinetics, and preliminary activity of the anti-IGF-1R antibody figitumumab (CP-751,871) in patients with sarcoma and Ewing's sarcoma: a phase I expansion cohort study*. *Lancet Oncol*, 2010. **11**(2): p. 129-35.
222. Juergens, H., et al., *Preliminary efficacy of the anti-insulin-like growth factor type I receptor antibody figitumumab in patients with refractory Ewing sarcoma*. *J Clin Oncol*, 2011. **29**(34): p. 4534-40.
223. Corral, J., et al., *Pathologic complete responses after chemotherapy plus figitumumab in stage IV non-small-cell lung cancer*. *J Clin Oncol*, 2011. **29**(7): p. e163-5.
224. Kim, S.Y., X. Wan, and L.J. Helman, *Targeting IGF-1R in the treatment of sarcomas: past, present and future*. *Bull Cancer*, 2009. **96**(7): p. E52-60.
225. Pollak, M., *The insulin and insulin-like growth factor receptor family in neoplasia: an update*. *Nat Rev Cancer*, 2012. **12**(3): p. 159-69.
226. Vasilcanu, D., et al., *The cyclolignan PPP induces activation loop-specific inhibition of tyrosine phosphorylation of the insulin-like growth factor-1 receptor. Link to the phosphatidyl inositol-3 kinase/Akt apoptotic pathway*. *Oncogene*, 2004. **23**(47): p. 7854-62.
227. Vasilcanu, R., et al., *Insulin-like growth factor type-I receptor-dependent phosphorylation of extracellular signal-regulated kinase 1/2 but not Akt (protein kinase B) can be induced by picropodophyllin*. *Mol Pharmacol*, 2008. **73**(3): p. 930-9.
228. Mulvihill, M.J., et al., *Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor*. *Future Med Chem*, 2009. **1**(6): p. 1153-71.
229. Ji, Q.S., et al., *A novel, potent, and selective insulin-like growth factor-I receptor kinase inhibitor blocks insulin-like growth factor-I receptor signaling in vitro and inhibits insulin-like growth factor-I receptor dependent tumor growth in vivo*. *Mol Cancer Ther*, 2007. **6**(8): p. 2158-67.
230. Haluska, P., et al., *In vitro and in vivo antitumor effects of the dual insulin-like growth factor-I/insulin receptor inhibitor, BMS-554417*. *Cancer Res*, 2006. **66**(1): p. 362-71.

231. Liu, T.J., et al., *Inhibition of both focal adhesion kinase and insulin-like growth factor-I receptor kinase suppresses glioma proliferation in vitro and in vivo*. Mol Cancer Ther, 2007. **6**(4): p. 1357-67.
232. Tanno, B., et al., *Down-regulation of insulin-like growth factor I receptor activity by NVP-AEW541 has an antitumor effect on neuroblastoma cells in vitro and in vivo*. Clin Cancer Res, 2006. **12**(22): p. 6772-80.
233. Manara, M.C., et al., *Preclinical in vivo study of new insulin-like growth factor-I receptor--specific inhibitor in Ewing's sarcoma*. Clin Cancer Res, 2007. **13**(4): p. 1322-30.
234. Gable, K.L., et al., *Diarylureas are small-molecule inhibitors of insulin-like growth factor I receptor signaling and breast cancer cell growth*. Mol Cancer Ther, 2006. **5**(4): p. 1079-86.
235. Economou, M.A., et al., *Inhibition of VEGF secretion and experimental choroidal neovascularization by picropodophyllin (PPP), an inhibitor of the insulin-like growth factor-I receptor*. Acta Ophthalmol, 2008. **86 Thesis 4**: p. 42-9.
236. Girnita, A., et al., *The insulin-like growth factor-I receptor inhibitor picropodophyllin causes tumor regression and attenuates mechanisms involved in invasion of uveal melanoma cells*. Acta Ophthalmol, 2008. **86 Thesis 4**: p. 26-34.
237. Ahlen, J., et al., *Insulin-like growth factor type I receptor expression correlates to good prognosis in highly malignant soft tissue sarcoma*. Clin Cancer Res, 2005. **11**(1): p. 206-16.
238. Economou, M.A., et al., *Oral picropodophyllin (PPP) is well tolerated in vivo and inhibits IGF-1R expression and growth of uveal melanoma*. Acta Ophthalmol, 2008. **86 Thesis 4**: p. 35-41.
239. Klinakis, A., et al., *Igf1r as a therapeutic target in a mouse model of basal-like breast cancer*. Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2359-64.
240. Nussbaum, T., et al., *Autocrine insulin-like growth factor-II stimulation of tumor cell migration is a progression step in human hepatocarcinogenesis*. Hepatology, 2008. **48**(1): p. 146-56.
241. Menu, E., et al., *Targeting the IGF-1R using picropodophyllin in the therapeutical 5T2MM mouse model of multiple myeloma: beneficial effects on tumor growth, angiogenesis, bone disease and survival*. Int J Cancer, 2007. **121**(8): p. 1857-61.
242. Razuvaev, A., et al., *The cyclolignan picropodophyllin attenuates intimal hyperplasia after rat carotid balloon injury by blocking insulin-like growth factor-I receptor signaling*. J Vasc Surg, 2007. **46**(1): p. 108-15.
243. Ekman, S., et al., *Clinical Phase I study with an Insulin-like Growth Factor-I receptor inhibitor: experiences in patients with squamous non-small cell lung carcinoma*. Acta Oncol, 2011. **50**(3): p. 441-7.
244. Gualberto, A. and M. Pollak, *Emerging role of insulin-like growth factor receptor inhibitors in oncology: early clinical trial results and future directions*. Oncogene, 2009. **28**(34): p. 3009-21.
245. Gao, J., et al., *Dual IGF-I/II-neutralizing antibody MEDI-573 potently inhibits IGF signaling and tumor growth*. Cancer Res, 2011. **71**(3): p. 1029-40.
246. Ren, S.G., et al., *Somatostatin analog induces insulin-like growth factor binding protein-1 (IGFBP-1) expression in human hepatoma cells*. Endocrinology, 1992. **131**(5): p. 2479-81.
247. Baserga, R., *The decline and fall of the IGF-I receptor*. J Cell Physiol, 2013. **228**(4): p. 675-9.

248. Kolb, H.C., M.G. Finn, and K.B. Sharpless, *Click Chemistry: Diverse Chemical Function from a Few Good Reactions*. *Angew Chem Int Ed Engl*, 2001. **40**(11): p. 2004-2021.
249. Tang, W. and M.L. Becker, "Click" reactions: a versatile toolbox for the synthesis of peptide-conjugates. *Chem Soc Rev*, 2014. **43**(20): p. 7013-39.
250. Tornøe, C.W. and M. Meldal, *Peptidotriazoles: Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions on Solid-Phase*, in *Peptides: The Wave of the Future: Proceedings of the Second International and the Seventeenth American Peptide Symposium, June 9–14, 2001, San Diego, California, U.S.A.*, M. Lebl and R.A. Houghten, Editors. 2001, Springer Netherlands: Dordrecht. p. 263-264.
251. Rostovtsev, V.V., et al., A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew Chem Int Ed Engl*, 2002. **41**(14): p. 2596-9.
252. Tornøe, C.W., C. Christensen, and M. Meldal, *Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides*. *J Org Chem*, 2002. **67**(9): p. 3057-64.
253. Worrell, B.T., J.A. Malik, and V.V. Fokin, *Direct evidence of a dinuclear copper intermediate in Cu(I)-catalyzed azide-alkyne cycloadditions*. *Science*, 2013. **340**(6131): p. 457-60.
254. Bock, V.D., H. Hiemstra, and J.H. van Maarseveen, *CuI-Catalyzed Alkyne–Azide "Click" Cycloadditions from a Mechanistic and Synthetic Perspective*. *European Journal of Organic Chemistry*, 2006. **2006**(1): p. 51-68.
255. Meldal, M. and C.W. Tornøe, *Cu-Catalyzed Azide–Alkyne Cycloaddition*. *Chemical Reviews*, 2008. **108**(8): p. 2952-3015.
256. Yoon, K., P. Goyal, and M. Weck, *Monofunctionalization of Dendrimers Using Microwave–Assisted 1,3 Dipolar Cycloadditions*. *Organic letters*, 2007. **9**(11): p. 2051-2054.
257. Miller, N., G.M. Williams, and M.A. Brimble, *Synthesis of fish antifreeze neoglycopeptides using microwave-assisted "click chemistry"*. *Org Lett*, 2009. **11**(11): p. 2409-12.
258. Chan, T.R., et al., *Polytriazoles as copper(I)-stabilizing ligands in catalysis*. *Org Lett*, 2004. **6**(17): p. 2853-5.
259. Hexnerova, R., et al., *Probing Receptor Specificity by Sampling the Conformational Space of the Insulin-like Growth Factor II C-domain*. *Journal of Biological Chemistry*, 2016. **291**(40): p. 21234-+.
260. Bright, G.M., *Recombinant IGF-I: Past, present and future*. *Growth Horm IGF Res*, 2016. **28**: p. 62-5.
261. Saegusa, J., et al., *The direct binding of insulin-like growth factor-1 (IGF-1) to integrin alphavbeta3 is involved in IGF-1 signaling*. *J Biol Chem*, 2009. **284**(36): p. 24106-14.
262. Whittaker, J., et al., *alpha-Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase*. *Proc Natl Acad Sci U S A*, 2012. **109**(28): p. 11166-71.
263. Liu, F., et al., *Chemical synthesis of peptides within the insulin superfamily*. *J Pept Sci*, 2016. **22**(5): p. 260-70.
264. Hossain, M.A., et al., *Solid phase synthesis and structural analysis of novel A-chain dicarba analogs of human relaxin-3 (INSL7) that exhibit full biological activity*. *Org Biomol Chem*, 2009. **7**(8): p. 1547-53.

265. Zhang, S., et al., *Role of the intra-A-chain disulfide bond of insulin-like peptide 3 in binding and activation of its receptor, RXFP2*. *Peptides*, 2010. **31**(9): p. 1730-1736.
266. Hossain, M.A., et al., *Chemically synthesized dicarba H2 relaxin analogues retain strong RXFP1 receptor activity but show an unexpected loss of in vitro serum stability*. *Org Biomol Chem*, 2015. **13**(44): p. 10895-903.
267. Shabanpoor, F., et al., *Minimization of human relaxin-3 leading to high-affinity analogues with increased selectivity for relaxin-family peptide 3 receptor (RXFP3) over RXFP1*. *J Med Chem*, 2012. **55**(4): p. 1671-81.
268. Wu, F., et al., *Synthesis of Four-Disulfide Insulin Analogs via Sequential Disulfide Bond Formation*. *J Org Chem*, 2017. **82**(7): p. 3506-3512.
269. Karas, J.A., et al., *Total Chemical Synthesis of an Intra-A-Chain Cystathionine Human Insulin Analogue with Enhanced Thermal Stability*. *Angew Chem Int Ed Engl*, 2016. **55**(47): p. 14743-14747.
270. Ruegg, U.T. and H.G. Gattner, *Reduction of S-sulpho groups by tributylphosphine: an improved method for the recombination of insulin chains*. *Hoppe Seylers Z Physiol Chem*, 1975. **356**(10): p. 1527-33.
271. Kosinová, L., et al., *Insight into the Structural and Biological Relevance of the T/R Transition of the N-Terminus of the B-Chain in Human Insulin*. *Biochemistry*, 2014. **53**(21): p. 3392-3402.
272. Křížková, K., et al., *Structural and Functional Study of the GlnB22-Insulin Mutant Responsible for Maturity-Onset Diabetes of the Young*. *PLoS ONE*, 2014. **9**(11): p. e112883.
273. Sohma, Y. and S.B. Kent, *Biomimetic synthesis of lispro insulin via a chemically synthesized "mini-proinsulin" prepared by oxime-forming ligation*. *J Am Chem Soc*, 2009. **131**(44): p. 16313-8.
274. Li, C.H., et al., *Total synthesis of insulin-like growth factor I (somatomedin C)*. *Proc Natl Acad Sci U S A*, 1983. **80**(8): p. 2216-20.
275. Iwai, M., et al., *Direct identification of disulfide bond linkages in human insulin-like growth factor I (IGF-I) by chemical synthesis*. *J Biochem*, 1989. **106**(6): p. 949-51.
276. Bagley, C.J., et al., *Synthesis of insulin-like growth factor I using N-methyl pyrrolidinone as the coupling solvent and trifluoromethane sulphonic acid cleavage from the resin*. *Int J Pept Protein Res*, 1990. **36**(4): p. 356-61.
277. Luisier, S., et al., *Total chemical synthesis of human proinsulin()*. *Chemical communications (Cambridge, England)*, 2010. **46**(43): p. 10.1039/c0cc03141k.
278. Sohma, Y., et al., *Design and folding of [GluA4(ObetaThrB30)]insulin ("ester insulin"): a minimal proinsulin surrogate that can be chemically converted into human insulin*. *Angew Chem Int Ed Engl*, 2010. **49**(32): p. 5489-93.
279. Machackova, K., et al., *Insulin-like Growth Factor I Analogs Clicked in the C Domain: Chemical Synthesis and Biological Activities*. *J Med Chem*, 2017. **60**(24): p. 10105-10117.
280. Behrendt, R., P. White, and J. Offer, *Advances in Fmoc solid-phase peptide synthesis*. *J Pept Sci*, 2016. **22**(1): p. 4-27.
281. Vikova, J., et al., *Rational steering of insulin binding specificity by intra-chain chemical crosslinking*. *Sci Rep*, 2016. **6**: p. 19431.
282. Isidro-Llobet, A., M. Alvarez, and F. Albericio, *Amino acid-protecting groups*. *Chem Rev*, 2009. **109**(6): p. 2455-504.
283. Hunter, M.J. and E.A. Komives, *Deprotection of S-acetamidomethyl cysteine-containing peptides by silver trifluoromethanesulfonate avoids the oxidation of methionines*. *Anal Biochem*, 1995. **228**(1): p. 173-7.

284. Fujita, M., et al., *An integrin binding-defective mutant of insulin-like growth factor-1 (R36E/R37E IGF1) acts as a dominant-negative antagonist of the IGF1 receptor (IGF1R) and suppresses tumorigenesis but still binds to IGF1R*. J Biol Chem, 2013. **288**(27): p. 19593-603.
285. Johnson, I.S., *Human insulin from recombinant DNA technology*. Science, 1983. **219**(4585): p. 632-7.
286. Zhou, P. and G. Wagner, *Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies*. J Biomol NMR, 2010. **46**(1): p. 23-31.
287. Tao, Y., et al., *Mechanisms of disease: signaling of the insulin-like growth factor 1 receptor pathway--therapeutic perspectives in cancer*. Nat Clin Pract Oncol, 2007. **4**(10): p. 591-602.
288. Schaffer, L., et al., *Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks*. Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4435-9.
289. Kaarsholm, N.C., et al., *Engineering stability of the insulin monomer fold with application to structure-activity relationships*. Biochemistry, 1993. **32**(40): p. 10773-8.
290. Weiss, M.A., et al., *Activities of monomeric insulin analogs at position A8 are uncorrelated with their thermodynamic stabilities*. J Biol Chem, 2001. **276**(43): p. 40018-24.
291. Wan, Z., et al., *Enhancing the activity of insulin at the receptor interface: crystal structure and photo-cross-linking of A8 analogues*. Biochemistry, 2004. **43**(51): p. 16119-33.
292. Conlon, J.M., *Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships*. Peptides, 2001. **22**(7): p. 1183-1193.
293. Simon, J., et al., *Enhanced Binding Affinity of Chicken Insulin in Rat-Liver Membranes and Human Lymphocytes - Relationship to Kinetic-Properties of Hormone-Receptor Interaction*. Endocrinology, 1977. **100**(1): p. 115-121.
294. Herring, R., R.H. Jones, and D.L. Russell-Jones, *Hepatoselectivity and the evolution of insulin*. Diabetes Obesity & Metabolism, 2014. **16**(1): p. 1-8.