Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
ELSEVIER
Bioorganic & Medicinal Chemistry Letters 14 (2004) 1407–1410

Structure-based de novo design of ligands using a three-dimensional model of the insulin receptor☆

Christopher Tan,[a,b] Lianhu Wei,[a,c] F. Peter Ottensmeyer,[d] Ira Goldfine,[f] Betty A. Maddux,[f] Cecil C. Yip,[e] Robert A. Batey[a,b,*] and Lakshmi P. Kotra[a,b,c,*]

[a]*Molecular Design and Information Technology Center, University of Toronto, 19 Russell Street, Toronto, Ontario, Canada*
[b]*Davenport Research Laboratories, Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario, Canada*
[c]*Department of Pharmaceutical Sciences, Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada*
[d]*Ontario Cancer Institute, Princess Margaret Hospital, 620 University Ave., Toronto, Ontario, Canada*
[e]*Banting and Best Department of Medical Research, University of Toronto, Ontario, Canada*
[f]*University of California at San Francisco, San Francisco, California, USA*

Received 7 September 2003; revised 22 January 2004; accepted 22 January 2004

Abstract—For the first time, a three-dimensional model of the insulin receptor is used in the de novo design of novel ligands that potentially mimic interactions of insulin at its receptor. Compound **4** competed with insulin as seen in autophosphorylation assays and inhibited up to 68% of IR autophosphorylation at 300 μM of **4** in 3T3IR cells induced by 1 nM insulin. This model provides a basis for the design of potent insulin receptor ligands.

A number of metabolic diseases, including diabetes which is reaching epidemic proportion in North America, are the results of abnormal interactions between insulin and its receptor.[1] Since insulin has to be administered by injection in the treatment of these diseases, the development of ingestable insulin mimetics has been of great interest in recent years. Several attempts were made to design molecules that can affect the insulin receptor, either its extracellular ligand-binding sites or at its intracellular tyrosine kinase domain.[1,2] Past attempts used either crystallographic information available on insulin, its mutational data, or the X-ray crystal structures of the intracellular components of the insulin receptor. Due to the absence of the structural information on the interactions of insulin at its receptor such as an X-ray crystal structure, a structure-based drug design against this receptor has not been possible. Additionally, such a design process against receptors is not as trivial as it is against an enzyme due to the complexities involved in the receptor signaling processes. Here, we disclose the refinement of a three-dimensional model of insulin–insulin receptor complex and its use in an insulin receptor-based de novo strategy to design molecules that interfere with the interactions of insulin at its receptor site.

The quaternary structure of the complex of insulin–insulin receptor was recently constructed using electron cryomicroscopy (EM).[3,4] An atomic model was built by fitting the X-ray crystallographic and NMR structures of various domain substructures of the insulin receptor and that of insulin into the above EM structure.[3] Through this exercise, a contiguous and complete three-dimensional model of complex of the insulin–insulin receptor was derived.

The insulin-binding site on the insulin receptor was identified in the electron microscopic reconstruction using gold-labeled insulin, and the X-ray crystal structure of insulin was docked into the binding pocket in such a way that various interactions between the amino acids in the receptor and the amino acids in insulin are complementary.[3,4] Herein we describe the refinement of a model of the insulin–insulin receptor complex, de novo design of ligands, their synthesis and their ability to bind the insulin-binding site.

Keywords: Insulin receptor; De novo ligand deign.
☆Supplementary data associated with this article can be found, in the online version at, doi: 10.1016/j.bmcl.2004.01.064
* Corresponding authors. Tel.: +1-416-978-5058 (R.A.B.); tel.: +1-416-978-2882; fax: +1-416-978-8511 (L.P.K.); e-mail: rbatey@chem.utoronto.ca; p.kotra@utoronto.ca.

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.01.064

Initially, the three-dimensional model of the extracellular portion of the insulin–insulin receptor complex was refined including disulfide bonds for computer modeling experiments. This model was subjected to energy-minimization, and the interactions between insulin and the insulin receptor were dissected.[5] This three-dimensional model was considered as a valid starting point not only to understand the interactions between insulin and its receptor, but also as a template pharmacophore to design nonpeptidic insulin receptor ligands. An analysis of the Connolly surface complemented by individual amino acid characteristics in the insulin-binding pocket revealed noticeable interactions between the insulin receptor and its ligand, insulin, that could be exploited for drug design. It is noteworthy here that although insulin receptor is a homodimer, the interactions between the insulin receptor and insulin are not symmetric. The obvious challenge is how to develop small molecule ligands that bind to the insulin receptor, for which the endogenous ligand, insulin, has a large surface of interactions.[4]

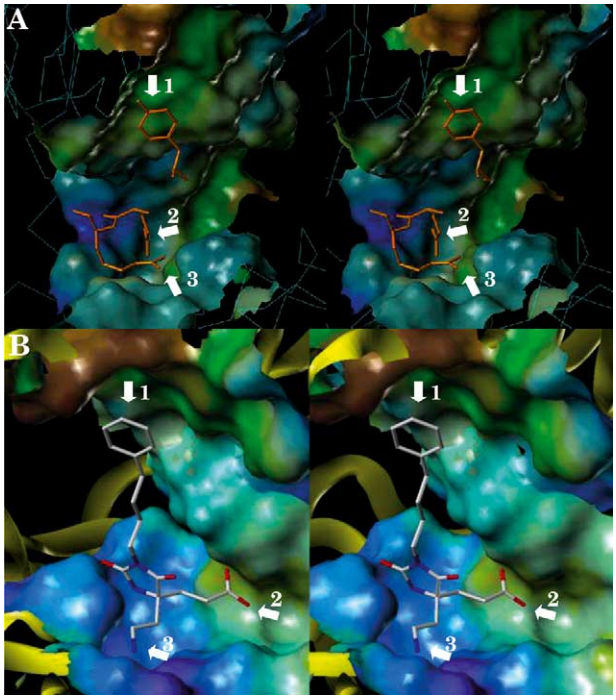
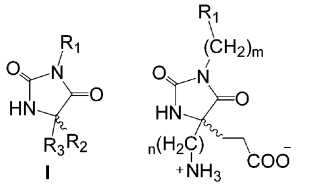


Figure 1. (A) Stereo view of the Connolly water accessible surface at the binding pocket of insulin receptor. The lipophilic character of the amino acid residues is mapped on to the surface (hydrophilic in blue, and hydrophobic in dark green to brown). The amino acid residues from insulin, TyrB16, GluB21 and ArgB22 (at the identified hydrophobic, cationic and anionic regions, respectively) are shown by the block arrows 1, 2 and 3, respectively. The backbone of the insulin receptor is shown by line representation in cyan and the three residues from insulin are shown as capped sticks in orange. (B) Stereo view of the energy-minimized model of a complex of compound **4** bound in the binding pocket of the refined model of the insulin receptor. Compound **4** is represented by a capped-sticks model color coded according to atom types (C: white, N: blue, O: red). Region 1 of the Connolly surface is shown by hydrophilic potential properties (brown: hydrophobic, blue: hydrophilic) and surface on the regions 2 and 3 are colored by electrostatic potential properties (purple-blue: electronegative, red-brown: electropositive).

Initially, relevant hydrophobic and ionic pockets on the extracellular domain of the insulin receptor were identified to characterize the binding site(s) for ligand design. These regions were contiguously mapped to create the binding pocket on the monomer 1 (regions 1 and 2 in Fig. 1A) spanning towards monomer 2 (region 3). This region of the binding pocket has three residues from insulin, TyrB16, GluB21 and ArgB22, which were seen as 'direct interaction regions' with the insulin receptor (Fig. 1A).

Following the above criteria, core structure **I** (Chart 1) was selected from our compound repository that can be readily derivatized at various positions. This core structure can carry three substitutions on the hydantoin ring providing an ability to attach three different 'functional groups' that can confer hydrophobic, cationic or anionic interactions at the binding regions, as needed. The lead template was docked into the binding site of the insulin receptor, from which compound **4** emerged as a representative ligand with appropriate side-chain substitutions extending into the hydrophobic (region 1, Fig. 1A), cationic (region 2) and anionic (region 3) pockets at the binding site. Specifically, compound **4** was docked into the binding pocket of insulin receptor using the Sybyl software program.[6] The three side chains of compound **4** docked into the binding pocket of the insulin receptor were positioned to mimic the interaction sites of insulin residues TyrB16, GluB21 and ArgB22. This docked complex was then subjected to energy minimization to obtain a reasonable model of the complex of compound **4** with the insulin receptor (Fig. 1B).[7]

Based on this model, a series of compounds **1**–**9** were designed and synthesized to understand the structure activity relationships. This series of compounds was designed to investigate the fidelity of the identified binding sites, by variation of the side-chain lengths attached to the hydantoin core, and the role of phenyl group versus a naphthyl group, to arrive at a reasonable lead compound. The structural diversity in the above set of nine compounds also spans the conformational space to overcome the lack of information on the dynamic nature of the receptor-binding site, although at the expense of introducing entropic penalties into the molecules.



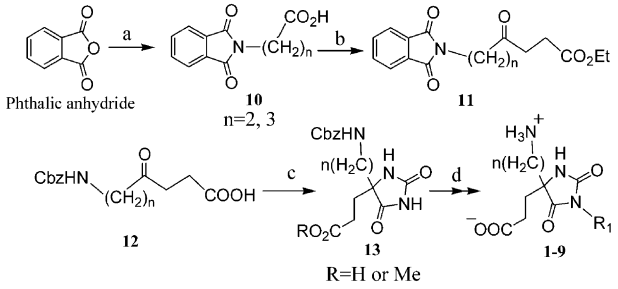
	R_1	m	n
1	Phenyl	1	2
2	Phenyl	2	2
3	Phenyl	3	2
4	Phenyl	4	2
5	Phenyl	5	2
6	Phenyl	6	2
7	Phenyl	3	3
8	Phenyl	4	3
9	Naphthyl	2	2

Chart 1.

Phthalimides **10** ($n=2$ or 3) were prepared by the addition of β-alanine or γ-aminobutyric acid to phthalic anhydride (Scheme 1). Compounds **10** were first converted to their respective acid chlorides and then, subjected to a palladium-catalyzed coupling reaction with a zinc homoenolate to obtain 6-phthalimido-4-oxo-hexanoic acid ethyl ester or 7-phthalimido-4-oxo-heptanoic acid ethyl ester **11** ($n=2$ or 3, respectively).[8,9] Phthalyl and ethyl groups on compounds **11** were hydrolyzed and the amine groups were protected as benzyl carbamates to yield **12**.[10] Compounds **12** were subjected to Bucherer–Bergs hydantoin formation to yield **13** (R=H, $n=2$ or 3).[11] Interestingly, when the free acid **12** was protected as a benzyl ester, hydantoin formation was unsuccessful, instead resulting mainly in amidation of the benzyl ester. The carboxylic acid functional group in **13** (R=H) was protected as a methyl ester **13** (R=Me) and subjected to alkylation at the N-3 position using Mitsunobu conditions[12] to yield the corresponding fully protected hydantoin derivatives. Deprotection of the Cbz protecting groups then afforded the target hydantoin derivatives **1–9** as their hydrochloride salts.

Whilst we recognized the high risk nature of the model built using the EM topological mapping fit with X-ray crystal structures of individual domains, and the uncertainties associated with structure-based design of molecules against receptors, the initial design process based on the model of the complex of insulin and the insulin receptor was expected to provide a lead-like structure. It was also obvious to us that the discovery of an agonist or an antagonist will need a better understanding of the structure–activity relationship study based on this three-dimensional model. Moreover, due to the large size of this membrane receptor, we utilized only the extracellular region of the insulin receptor containing the insulin interacting domains.

With the above premises in mind, compounds **1–9** were evaluated for their autophosphorylation activity, at various concentrations of insulin, in a cell-based assay using 3T3IR cells expressing insulin receptor.[13] If the designed compounds were agonists, then one would see a synergistic activity with insulin, whilst if they are simply binding at the insulin-binding site, the autophosphorylation activity would be lowered compared to that using insulin alone. Our primary focus was on the



Scheme 1. Reagents. (a) AcOH, appropriate amino acid, Δ; (b) $SOCl_2$, Δ and then, $Zn(CH_2CH_2CO_2Et)_2$, $Pd(PPh_3)_2Cl_2$, C_6H_6/Et_2O; (c) $HCl/HCO_2H/H_2O$, Δ followed by $ClCO_2CH_2C_6H_5$, K_2CO_3, Et_2O, H_2O, followed by $(NH_4)_2CO_3$, KCN, EtOH, H_2O, Δ; (d) four steps, see Supporting Information.

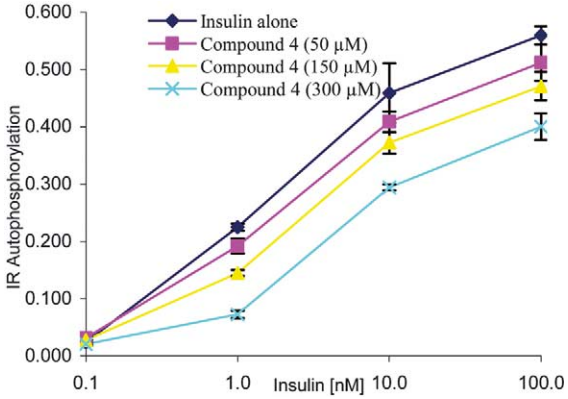


Figure 2. Evaluation of binding ability of compound **4** in competition with insulin (in vivo autophosphorylation of insulin stimulated 3T3IR cells±2 h incubation with compound **4**).

'binding ability' of these compounds to the insulin receptor as seen by the decrease in the phosphorylation of insulin receptor when activated with insulin, its natural ligand. Compound **4** showed an observable and significant binding to the insulin receptor as concluded from its ability to decrease the autophosphorylation of insulin receptor in the presence of insulin, at concentrations of **4** from 50 μM to 300 μM (Fig. 2). At a concentration of 150 μM of compound **4**, insulin-induced autophosphorylation levels were inhibited by 36% and 19% at 1 nM and 10 nM of insulin concentrations, respectively. At a higher concentration of 300 μM of compound **4**, insulin-induced autophosphorylation was inhibited by 68% and 36% at 1 nM and 10 nM concentrations of insulin, respectively. This indicates that compound **4** may be interacting or competing with insulin at the insulin binding site and inhibiting agonistic activity of insulin. Other compounds **1–3** and **5–9** did not show observable and/or significant effects up to 250 μM concentration (data not shown here). All the tested compounds have different side chain lengths connecting the functional moieties, and the differences in these side chains presumably do not permit favorable structural interactions to allow binding at the insulin receptor-binding pocket. The conformational flexibilities introduced into the methylene groups of the three substitutions on the hydantoin core in compound **4** may have resulted in modest potency of compound **4**.

In conclusion, compound **4**, designed using an insulin-insulin receptor model, exhibited optimal chain lengths amongst a set of designed compounds based upon a hydantoin core structure, antagonizing the binding of insulin. While it is necessary to optimize the structures of such binders to the insulin receptor, the above results show that the model of the complex of insulin receptor-insulin is a valid starting point for structure-based drug discovery efforts to design insulin mimetics.

Acknowledgements

This research was supported by a grant from the Canadian Institutes of Health Research-Proof of Principle program and by a grant from the American

Diabetes Association. L.P.K. is a recipient of Rx&D HRF-CIHR research career award. R.A.B gratefully acknowledges receipt of a Premier's Research Excellence Award. Infrastructure grant from Ontario Innovation Trust for the Molecular Design and Information Technology Center is acknowledged.

References and notes

1. (a) Balasubramanyam, M.; Mohan, V. *J. Biosci*. **2001**, *26*, 383. (b) Nuss, J. M.; Wagman, A. S. *Ann. Rep. Med. Chem*. **2000**, *35*, 211.
2. (a) Liu, K.; Xu, L.; Szalkowiski, D.; Li, Z.; Ding, V.; Kwei, G.; Huskey, S.; Moller, D. E.; Heck, J. V.; Zhang, B. B.; Jones, A. B. *J. Med. Chem*. **2000**, *43*, 3487. (b) Schlein, M.; Ludvigsen, S.; Olsen, H. B.; Andersen, A. S.; Danielsen, G. M.; Kaarshlom, N. C. *Biochemistry* **2001**, *40*, 13520 and references cited therein.
3. Luo, R. Z.; Beniac, D. R.; Fernandes, A.; Yip, C. C.; Ottensmeyer, F. P. *Science* **1999**, *285*, 1077.
4. (a) Ottensmeyer, F. P.; Beniac, D. R.; Luo, R. Z.; Yip, C. C. *Biochemistry* **2000**, *39*, 12103. (b) Yip, C. C.; Ottensmeyer, F. P. *J. Biol. Chem*. **2003**, *278*, 27329.
5. Energy minimizations were carried out using Amber version 6 suite of programs (UCSF). Visualization, three-dimensional structural analysis and design were carried out using Sybyl version 6.7 suite of programs (Tripos, Inc., USA). The surface construction was carried out by MOLCAD module available in Sybyl suite of programs.
6. The docking exercise was performed either manually or using Sybyl suite of software using the 'DOCK' procedure available as part of this software. First, compound **4** was constructed and was energy-minimized using default parameters in Sybyl. This energy-refined structure was then used to dock into the binding pocket on the insulin receptor.
7. The electrostatic potential charges, bond lengths and angles for **4** were obtained using Gaussian98 density functional theory method (B3LYP/6-31G*). The complex of **4** and insulin receptor was immersed in a box of TIP3P water with 8 Å thickness from the surface of the protein. A total of 75,240 water molecules were added during solvation. During the energy minimization, the cutoff distance for non-bonded interactions was set at 12.0 Å.
8. (a) Wang, J.; Scott, I. *Tetrahedron Lett*. **1997**, *38*, 739. (b) Rasolonjatovo, I.; Sarfati, S. R. *Nucleosides & Nucleotides* **1998**, *19*, 2021.
9. Nakamura, E.; Aoki, S.; Sekiya, K.; Oshino, H.; Kuwajima, I. *J. Am. Chem. Soc*. **1987**, *109*, 8056.
10. Kim, S.; Lee, I. I.; Kim, Y. C. *J. Org. Chem*. **1985**, *50*, 560.
11. (a) Bergs, H. Ger. Pat. 566,094, 1929. (b) Bucherer, H. R.; Barsch, H. *J. Prakt. Chem*. **1934**, *140*, 151. (c) Abshire, C. J.; Planet, G. *J. Med. Chem*. **1972**, *15*, 226.
12. Mitsunobu, O. *Synthesis* **1981**, 1.
13. The ligand binding abilities of compounds **1**–**9** were carried out in in vivo cell systems using the following procedure: 3T3HIR Cells were grown in 96-well plates for 24 h, then media was changed to DMEM without serum. After 18 h, media was removed and the test compounds at various concentrations were added to wells at 37 °C and the cells were incubated for 2 h. Next, insulin (0–100 nM in DMEM, 0.1% BSA) was added and incubated for 10 min. After incubation, cells were washed and solubilized in 200 μL solubilization buffer (50 mM Hepes, pH 7.6, 150 mM NaCl, 1% Triton X-100, 1 mM PMSF). Lysate (100 μL) was added to wells coated with MA-20, and allowed to bind for 2 h. After washing, anti-phosphotyrosine antibody-horseradish peroxidase conjugate (PY20HRP, Santa Cruz Biotechnology) was added for 2 h, then washed again. Signal was detected using TMB peroxidase substrate (3,3′,5,5′-tetramethylbenzidine and H_2O2 in a citric acid buffer, KPL, Gaithersburg, MD). Several concentrations of the test compounds were considered starting from 1 μM to 300 μM and the experiments were performed in triplicate. The autophosphorylation was measured by optical density measurements at 451 nm. A plot of the optical density (ΔOD, difference between the sample and the blank reflecting IR autophosphorylation) versus various insulin concentrations (in nM) used for stimulation of 3T3 IR cells is constructed to analyze the binding ability of each test compound to the insulin receptor. Each data point is an average of the triplicate values ±SD.