Research Article

Journal of PeptideScience

Received: 25 April 2013 Revised: 18 June 2013 Accepted: 18 June 2013 Published online in Wiley Online Library: 26 July 2013

(wileyonlinelibrary.com) DOI 10.1002/psc.2538

A novel glucagon-like peptide 1 peptide identified from *Ophisaurus harti*

Jingjing Zhu,[a] Xian Huang,[b] Hong Gao,[b] Qiuying Bao,[b] Yun Zhao,[a] Jin-Feng Hu[c]* and Gang Xia[a,b]*

ABSTRACT: Glucagon-like peptide 1 receptor (GLP1R) is a promising target for the treatment of type 2 diabetes. Because of the short half-life of endogenous GLP1 peptide, other GLP1R agonists are considered to be appealing therapeutic candidates. A high-throughput assay has been established to screen for GLP1R agonists in a 60 000-well natural product compound library fractionated from 670 different herbs/materials widely used in traditional Chinese medicines (TCMs). The screening is based on primary screen of GLP1R$^+$ reporter gene assay with the counter screen in GLP1R$^-$ cell line. An active fraction, A089-147, was identified from the screening. Fraction A089-147 was isolated from dried *Ophisaurus harti*, and the fact that its GLP1R agonist activity was sensitive to trypsin treatment indicates its peptidic nature. The active ingredient of A089-147 was later identified as *O. harti* GLP1 through transcriptome analysis. Chemically synthesized *O. harti* GLP1 showed GLP1R agonist activity and sensitivity to dipeptidase IV digestion. This study illustrated a comprehensive screening strategy to identify novel GLP1R agonists from TCMs libraries and at the same time underlined the difficulty of identifying a non-peptidic GLP1R agonist. The novel *O. harti* GLP1 peptide yielded from this study confirmed broader application of TCMs libraries in active peptide identification. Copyright © 2013 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: GLP1; GLP1R; High throughput screening; type 2 diabetes; TCM; transcriptome; *Ophisaurus harti*

Introduction

Type 2 diabetes is a chronic, age-related degenerative disorder that is the leading cause of morbidity and mortality in the older people [1]. Glucagon-like peptide 1 receptor (GLP1R) has been a target for treatment of type 2 diabetes[2]. GLP1R can be activated *in vivo* by binding with GLP1, leading to an increase in glucose-dependent insulin gene expression and insulin secretion [3]. Various studies have demonstrated that GLP1 is effective in lowering glucose level without the risk of hypoglycemia associated with insulin overdose[4,5].

Endogenously produced GLP1 peptide can be rapidly inactivated by a serum protease, dipeptidase IV (DPP-IV), resulting in a half-life of only 2 min [6]. Hence, tremendous efforts have been made in search of molecules that prolong the activation of GLP1R[7]. These include GLP1 mimetics with resistance to DPP-IV cleavage, small molecule GLP1R agonists, or DPP-IV inhibitors [8,9] that indirectly increase the half-life of endogenous GLP1. A successful example is exenatide, a peptide originally isolated in the venom of the Gila monster (*Heloderma suspectum*)[10]. Exenatide is a potent GLP1R agonist with multiple glucoregulatory effects for the treatment of type 2 diabetes [11–13]. Exenatide has superior pharmacokinetics when compared with the endogenous GLP1 – resistance of exenatide to DPP-IV cleavage extends its biological half-life to 30 min *in vivo* [14].

In this study, a natural product library containing 60 000 fractions extracted from 671 herbs/materials was screened to identify the novel GLP1R agonist. These herbs/materials are routinely used in traditional Chinese medicines (TCMs). In one of the earliest books about Chinese medical theory - the 'Huangdi Neijing', compiled around 100 BC - the condition known as 'xiaoke' is mentioned. This is translated today as diabetes or diabetic exhaustion[15]. By the latter half of the 20th century, about 200 standard prescriptions were recorded for treating diabetes. Even in the current era, many studies demonstrate that TCM is efficacious in treating type 2 diabetes [16–18]. The goal of this study was to utilize current drug discovery technologies to identify and characterize novel GLP1R agonists from this natural product library.

Materials and Methods

Cell Culture and Stable Cell Lines Establishment

The wild type human embryonic kidney 293 (HEK293) cells were cultured in Dulbecco's Modified Eagle Medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (Invitrogen,

* *Correspondence to: Gang Xia, Key Laboratory of Brain Functional Genomics, Ministry of Education, East China Normal University (ECNU), Shanghai 200062, China. E-mail: gxia@brain.ecnu.edu.cn or gang.xia@gmail.com Jin-Feng Hu, Department of Natural Products Chemistry, School of Pharmacy, Fudan University, 826 Zhangheng Road, Shanghai 201203, China. E-mail: jfhu@fudan.edu.cn*

a *Institutes for Advanced Interdisciplinary Research, East China Normal University, 3663 N. Zhongshan Road, Shanghai 200062, China*

b *Key Laboratory of Brain Functional Genomics, Ministry of Education, East China Normal University, 3663 N. Zhongshan Road, Shanghai 200062, China*

c *Department of Natural Products Chemistry, School of Pharmacy, Fudan University, 826 Zhangheng Road, Shanghai 201203, China*

Abbreviations: *DMSO, dimethyl sulfoxide; DPP-IV, dipeptidase IV; FBS, fetal bovine serum; GLP1, glucagon-like peptide 1; GLP1R, glucagon-like peptide 1 receptor; HPLC, High Performance Liquid Chromatography; HTPS, high throughput screening; MWCO, molecular weight cutoff; O. harti, Ophisaurus harti; PBS, phosphate buffer saline; TCM, Traditional Chinese Medicine.*

Journal of PeptideScience

Carlsbad, CA, USA), 100 units/ml of penicillin sodium (Invitrogen, Carlsbad, CA, USA) and 100 units/ml of streptomycin sulfate (Invitrogen, Carlsbad, CA, USA) at 37 °C with 5% CO_2 in a humidified incubator. Stable GLP1R$^+$ cell colonies were selected from wild type HEK293 cells (Cell Bank, Chinese Academy of Sciences, Shanghai, China) transfected with pcDNA3.1(+)/GLP1R plasmids and pGL3/3MRE-CRE-luc plasmids (both plasmids were kindly provided by Dr Y. H. Hu [19,20]). A reference GLP1R$^-$ cell line was selected from HEK293 cells only transfected with pGL3/3MRE-CRE-luc plasmid. These two cell lines were cultured in a selection medium containing 500 μg/ml of G418 (Invitrogen, Carlsbad, CA, USA). Induced cell clones that gave a greater than 15-fold response in comparison with uninduced cells were selected for use in subsequent studies.

Compound Library

The natural product compound libraries screened in this study included 10 000 flash fractions and 50 000 high-performance liquid chromatography (HPLC) fractions that were generated from 671 different herbs/materials widely used in TCMs. This was accomplished by using protocols similar to previous works performed by one of the authors (J.-F. Hu) [21,22].

In brief, the flash library fractionation process proceeded as follows: each biomass (100 g) was extracted with ethanol/ethyl acetate (50 : 50). This extract was filtered and the organic layer was removed on a rotary evaporator under reduced pressure to yield a viscous mass. Then a two-gram aliquot of the mass was subjected to a silica gel flash column chromatography on Flash Master II (Argonant, USA) that was eluted by the stepwise gradients illustrated in Table 1. Two fractions were collected for each gradient step, and overall 20 fractions were collected for each herb. Flash fractions were dried and dissolved in dimethyl sulfoxide (DMSO) at a consistent concentration of 10 mg/ml.

The HPLC library fractionate process proceeded as follows: using the same extraction procedure mentioned in the previous texts, a one-gram aliquot of the ethanol/ethyl acetate (50 : 50) extracted mass was fractionated on silica columns by using the Flash Master II system with a step gradient consisting of gradients 2, 3, 5, 6, and 7 in Table 1 to obtain five flash fractions. Each flash fraction was further fractionated using two parallel semi-preparative HPLC systems (running at 3.5 ml/min and allowing 2 min per collection per tube) to yield 60 fractions for adequate separation. The HPLC consists of a Beckman Coulter System Gold 508 autosampler (Beckman Coulter, Brea, CA, USA), Gold 126 gradient HPLC pumps (Beckman Coulter, Brea, CA, USA) with a Beckman System Gold 168 (Beckman Coulter, Brea, CA, USA), a Sedex 80 ELSD (SEDERE, Olivet, France) detector and a Beckman Coulter Ultrasphare ODS column (250 × 10 mm, dp = 5 μm). Each HPLC fraction was dried and dissolved in 200 μl of DMSO.

Large Scale Preparation of the Active Fraction from *Ophisaurus harti*

Several approaches have been explored for the large-scale isolation of the active fraction with GLP1R agonist activity. During fractionation, each fraction was monitored for agonist activity by using a GLP1R reporter gene assay. The final method optimized for efficiency and its ability to scale up activity is described in the succeeding texts.

Dried *Ophisaurus harti* (Tong-Jun-ge pharmacy at Chongqing, China) was grounded into powder for extraction. A mixture of 400 g powder and 2000 ml of ethanol/ethyl acetate (50 : 50) was heated to 75 °C to maintain a 'microboiling' status for 2 h by using a water bath reflux system at normal atmosphere pressure. Then the 2000 ml of solvent was replaced with a fresh batch and the heating process repeated. This 'microboiling' extraction process was repeated twice for a total of three extractions.

After the third extraction, the remaining solid residues were filtered and extracted with 3 l of an aqueous solvent mixture of methanol/water (70 : 30) by using a water bath reflux system at normal atmosphere pressure at 80 °C for 3 h. The 'aqueous' extraction was repeated three times. All of the 12 000 ml of methanol/water fractions was collected and concentrated using a rotary evaporator at 45 °C to remove methanol and water under reduced pressure. The final volume of concentrated methanol/water extract was about 100 ml. Even though two completely different extraction methods were used, for consistency purpose, the large-scale prepared active fraction was named A089-147, which was the same name for the original screening fraction with GLP1R agonist activity.

The protein concentration of large scale prepared A089-147 was measured using the BCA (bicinchoninic acid) method following the manufacturer's instructions (Tiangen, Shanghai, China).

GLP1R Reporter Gene Assays

Forty microliter of GLP1R$^+$ or GLP1R$^-$ cells were seeded in a 384-well plate at 2000 cells/well and incubated overnight for cell recovery. The next morning, 100 nl (nanoliter) of compounds were transferred into the 384-well cell plates by using a 384 × 100 nl pintool head (V&P Scientific, San Diego, CA, USA) controlled by a BioMek FX liquid handler (Beckman Coulter, Brea, CA, USA). After 4 h of incubation under normal tissue culture condition, 20 μl of Bright-Glo™ (Promega, Madison, WI, USA) was added using a liquid dispenser

Table 1. Stepwise gradients used in generating flash fraction library

Step gradient #		1	2	3	4	5	6	7	8	9	10
Solvent (%)	Hexane	100	75	50	25	—	—	—	—	—	—
	Ethyl acetate	—	25	50	75	100	75	50	25	—	—
	Methanol	—	—	—	—	—	25	50	75	100	—
	Water	—	—	—	—	—	—	—	—	—	100

(MicroFill™, BioTek, Winooski, VT, USA). The luminescence signals were read on a multi-function plate reader (SpectraMax M5, Molecular Devices, Sunnyvale, CA, USA).

High-Throughput Identification and Confirmation of GLP1R Agonists

Following the primary screen using $GLP1R^+$ reporter gene assay to screen for GLP1R agonists from a 60 000-well natural product library; a counter screen to identify false positive hits using $GLP1R^-$ cells was conducted in parallel. Compounds showing activity in $GLP1R^-$ cells where the activation was independent of GLP1R were considered to be false positives.

Cyclic Adenosine Monophosphate Assay

In brief, $GLP1R^+$ cells were seeded at 5000 cells/well in 384-well plates and incubated overnight. The next morning, cells were incubated with testing samples for 2 h, then lysed to measure cellular cAMP levels by using cAMP HiRange Assay Kit (Cisbio US, Bedford, MA, USA) following manufacturer's instruction. The homogeneous time resolved fluorescence signals produced in this assay were measured on a SpectraMax M5 multi-function plate reader (Molecular Devices, Sunnyvale, CA, USA).

For trypsin treatment: 40 nM of GLP1, 4 nM of exenatide, or 20 μg/ml of A089-147 were incubated at 37 °C with 0.05 μg/ml of trypsin (Invitrogen, Carlsbad, CA, USA) in PBS/BSA buffer containing 1X PBS buffer (pH = 7.2) and 1 mg/ml BSA (Sigma-Aldrich, St. Louis, MO, USA). After 30 min of incubation, aliquots were saved and tested for their GLP1R agonist activity in reporter gene assays.

For DPP-IV treatment: 10 μM of GLP1, 10 μM of exenatide, or 10 μM of chemically synthesized *O. harti* GLP1 were incubated at 37 °C with 0.1 μg/ml of DPP-IV (Sigma-Aldrich, St. Louis, MO, USA) in PBS/BSA buffer. For samples with DPP-IV activity suppressed, a final concentration of 10 μM DPP-IV inhibitor KR-62436 [23] (Sigma-Aldrich, St. Louis, MO, USA) was added before incubation. After 30 min of incubation, aliquots were saved and tested for their GLP1R agonist activity in reporter gene assays.

Activity of Fresh Tissue Extracts From Different Organs of *O. harti*

Different tissues were obtained from dissecting fresh *O. harti* (Chongqing, China) and were homogenized. The homogenized extracts were centrifuged at 12 000 rpm for 5 min. Supernatants were normalized by protein concentration before being tested for GLP1R activation in a reporter gene assay.

Transcriptome Analysis of the Gastrointestinal Tract of *O. harti*

The total RNA from the gastrointestinal tract of *O. harti* was extracted using the QIAGEN RNeasy Maxi Kit (Qiagen, Hilden, Germany) and tested for quality using an Agilent 2100 Bioanalyzer (Amersham Biosciences, Uppsala, Sweden).

Transcriptome analysis was carried out by the Beijing Genomic Institute. In this study, 10 μg of total RNA isolated from *O. harti* was used in generating cDNA. Sequences were obtained using an Illumina HiSeq 2000 platform. The assembled *O. harti* GLP1 sequence was confirmed by sequence analysis of reverse transcription (RT-PCR) products using the same RNAs extracted from the gastrointestinal tract of *O. harti* as a template and primers (5'-CATGCTGCGGTTCACTTC-3' and 5'-GTTCCACAGCTCAGAGAC-3') based on the *in silico O. harti* GLP1 DNA sequence.

O. harti GLP1 Peptide Sources

Exenatide and human GLP1 peptide were purchased from the Chinese Peptide Company. (Hangzhou, China)

The *O. harti* GLP1 peptide was chemically synthesized by a peptide service company (Scilight Biotechnology, Beijing, China) using the Fmoc solid phase procedure and was HPLC purified. The final GLP1 peptide is more than 98% pure (OD220) with expected molecular weight ($M + H^+$ is 3956.78 vs theoretical M is 3955.60).

Results

Establishment and Validation of GLP1R Stable Cell Lines for High-Throughput Screen

Stable, sensitive cell lines responsive to ligand stimulation are the key to successful cell-based high-throughput screening. Twenty stable $GLP1R^+$ and $GLP1R^-$ cell clones were isolated and tested for their response to forskolin and GLP1 stimulation. $GLP1R^+$ and $GLP1R^-$ cell clones used in this study had at least a 20-fold increase in signaling upon stimulation [Figure 1(A)]. $GLP1R^-$ cells were used in counter-screens to identify false positive hits that activated the reporter gene assay in the absence of the GLP1 receptor.

GLP1R Agonist Screen and Confirmation

The primary screen to identify GLP1R agonists from the 60 000-well natural product library was conducted with $GLP1R^+$ cells. The activity distribution of the natural product library is shown in Figure 1(B). Fractions with agonist activity larger than five folds over background activity were considered as hits and subjected to the confirmation screen.

Both $GLP1R^+$ and $GLP1R^-$ cells were used in the confirmation screen. The established $GLP1R^+$ reporter gene assay for identifying agonist activity is highly reproducible, with a confirmation rate greater than 95%. However, all the fractions except for one showed positive activities in the $GLP1R^-$ cells, indicating that they were false positives. The only positive fraction (A089-147) from aqueous extracts of *O. harti* activated $GLP1R^+$, but not $GLP1R^-$ cells in a dose-dependent manner [Figure 1(C)].

GLP1R Agonist Activity Confirmation of Large Scale Prepared A089-147

Because of the large number of wells in the original library and the moderately complicated library preparation procedure, errors including mislabeled fractions or chromatograph column failures might have occurred during library preparation. An independent preparation of A089-147 from starting material was necessary to validate the activity of A089-147. The newly prepared A089-147 specifically activated $GLP1R^+$ cells with EC_{50} at 2–4 μg/ml [Figure 2 (A)], which is more potent than the original A089-147.

The GLP1 receptor mainly functions through the Gs pathway that uses elevated cAMP as a second messenger that in turn relays extracellular signals to the cell nucleus. The activity of A089-147 on cAMP is shown in Figure 2(B); A089-147 increased the concentration of cAMP is consistent with the presence of a GLP1R agonist in A089-147.

Journal of PeptideScience

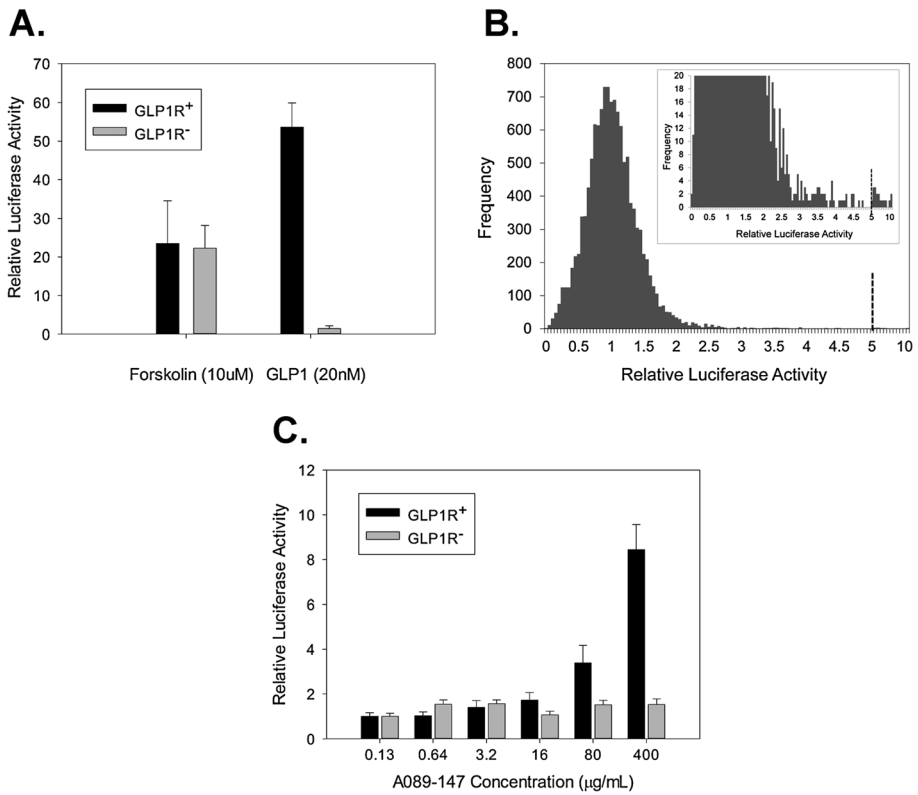

Figure 1. Results of high-throughput screening for glucagon-like peptide 1 receptor (GLP1R) agonists. (A) Responses of $GLP1R^{+}$ and $GLP1R^{-}$ cells upon stimulation with reference compound/peptide used in the screening. Concentrations shown are final concentrations incubated with cells. Relative luciferase activity was calculated as measured signals of treated cells divided by untreated cells. (B) Histogram of activity distribution in GLP1R agonist screening. The inserted graph is from the same result with an expanded *y*-axis. Library fractions with response window >5 (dotted line) were considered as 'hits'. (C) Dose-dependent confirmation of specific GLP1R agonist activity for fraction A089-147. A089-147 concentrations were calculated on the basis of library starting concentration of 10 mg/ml.

The Activity of A089-147 is from A Peptidic Macromolecule

In order to characterize the active ingredient in A089-147, fraction A089-147 was dialyzed in PBS buffer by using a 12 000–14 000 molecular weight cutoff membrane and tested in a GLP1R agonist reporter gene assay. Most of the A089-147 activity remained after the dialysis process. This indicated that the GLP1R agonist activity in A089-147 was not from a small molecule compound [Figure 3(A)].

Because of its large molecular weight, the active ingredient in A089-147 was considered to be a biological agent, e.g. protein or polynucleotides, and so on. A089-147 was treated with protease trypsin, which hydrolyzes proteins into small peptides thereby inactivating proteins, followed by a GLP1R reporter gene assay. Results from the trypsin treatment [Figure 3(B)] indicated that the active ingredient in A089-147 was likely a protein due to its susceptibility to trypsin treatment.

Gastrointestinal Extract from Fresh *O. harti* Activates GLP1R

Fraction A089-147 was extracted from the entire body of dried *O. harti*. In order to determine the origin of A089-147's active ingredient within *O. harti*, extracts from the organs of fresh *O. harti* were prepared and tested in GLP1R reporter gene assays. Extracts from the gastrointestinal tract had the highest activity compared with other organs (Figure 4). Because GLP1 from different species are secreted by intestinal L cells and concentrated in the intestinal region[24,25], the activity from the *O. harti* gastrointestinal tract was generated by a peptide homologous to GLP1.

GLP1 Sequence Identification and Confirmation

Transcriptome analysis has been adopted recently to study the functional elements of the genome and reveal the molecular constituents of cells and tissues [26,27]. Table 2 shows the transcriptome analysis results of the active ingredient. The potential genes of *O. harti* were annotated after a homologous search through five major protein/gene sequence databases (nucleotide, non-redundant, UniProt/Swiss-Prot, Kyoto encyclopedia of genes and genomes, and gene ontology). The sequence of assembled *O. harti* GLP1 was identical to the sequence of the RT-PCR product using primers based on the *in silico* GLP1 sequence. The homology relationship of GLP1 from different species is shown in Figure 5[28].

O. harti GLP1 is Sensitive to DPP-IV Cleavage

Endogenous GLP1 *in vivo* is rapidly inactivated by protease DPP-IV that cleaves X-Alanine/proline dipeptides from the *N*-terminus of polypeptides. Exenatide, that is secreted by the Gila monster, is resistant to DPP-IV. This leads to exenatide's development into a billion-dollar commercial treatment for type 2 diabetes. It is interested to know if *O. harti* GLP1, which is a close relative of the Gila monster [29], may also be resistant to DPP-IV cleavage. DPP-IV assay results indicate *O. harti* GLP1 has the similar sensitivity as human GLP1 (Figure 6).

Rapid inactivation by DPP-IV is consistent with the presence in *O. harti* GLP1 of the same *N*-terminal dipeptide His-Ala of human GLP1, despite their relatively low (46.7%) sequence homology.

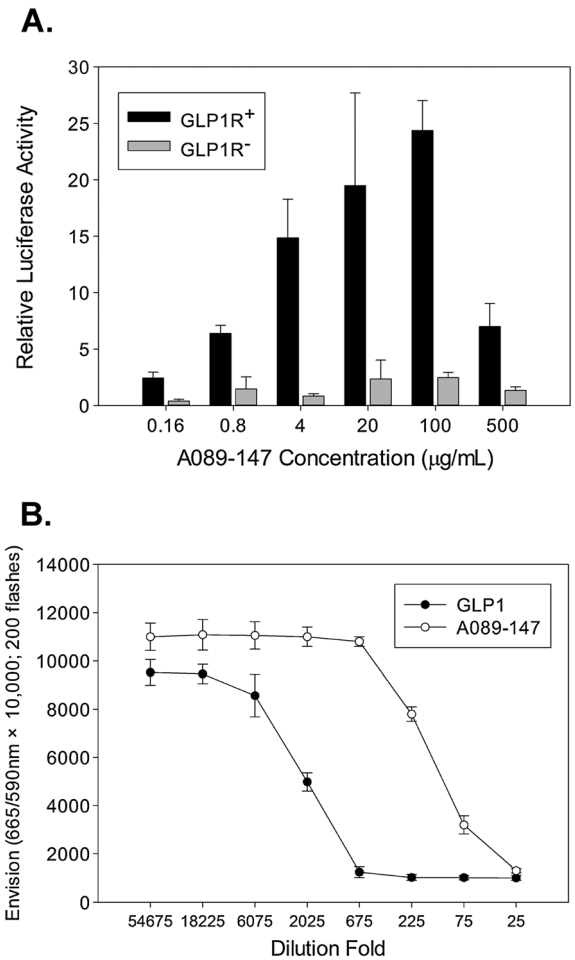

Figure 2. Confirmation of glucagon-like peptide 1 receptor (GLP1R) agonist activity from large scale prepared fraction A089-147. (A) Large scale prepared A089-147 activated GLP1R$^+$ cell, but not GLP1R$^-$ cell. Total protein concentration in fraction A089-147 is measured using the bicinchoninic acid method. (B) Cyclic adenosine monophosphate (cAMP) assay of human GLP1 and A089-147. The starting concentration of GLP1 is 40 nM, and the starting concentration of A089-147 protein is 20 μg/ml in the cAMP assay. Cyclic-AMP-d2 is a competitor of the endogenous cAMP in GLP1R$^+$ cells. Anti-cAMP-cryptate conjugates with cAMP-d2. As such, if the endogenous cAMP level is higher, then the signal of 665/590 nM is lower. All of the signal curves decreased as the concentrations of the sample increased, indicating that A089-147 was able to increase the cellular cAMP level in GLP1R$^+$ cells much like the GLP1 peptide.

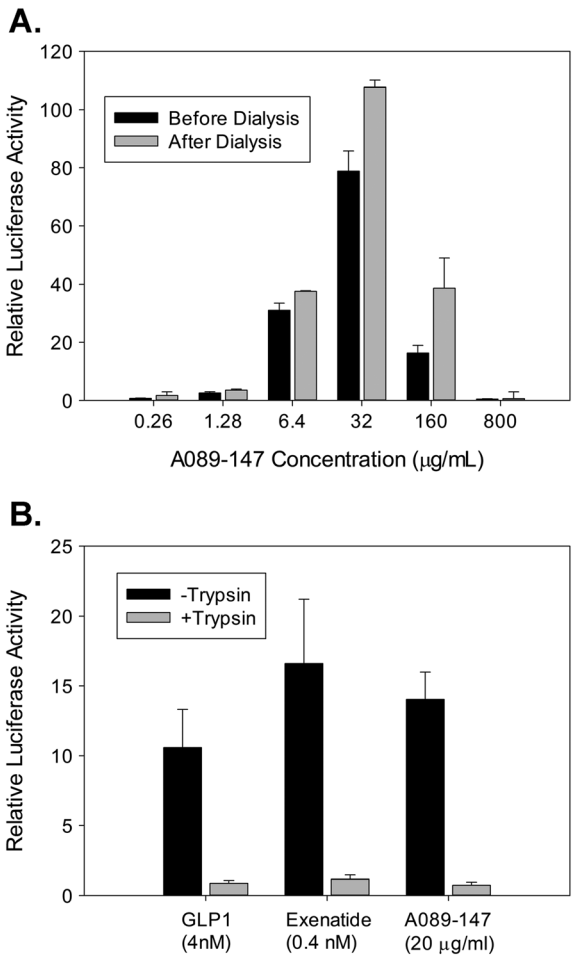

Figure 3. The activity of A089-147 is originated from a peptidic macromolecule. (A) The glucagon-like peptide 1 receptor agonist activity of A089-147 retained after dialysis though a dialysis membrane with 12 000 molecular weight cutoff. The protein concentration of A089-147 is measured using the bicinchoninic acid method. (B) Trypsin abolished the glucagon-like peptide 1 receptor agonist activity of fraction A089-147.

Discussion

A non-peptidic, orally active GLP1R agonist is considered to be a valuable treatment for type 2 diabetes. Despite extraordinary efforts to identify such a molecule by most major pharmaceutical companies, few successes have been achieved. Natural products have served as important starting points in drug discovery [30]; 63% of all small molecule drugs were naturally derived or semi-synthetic derivatives of natural products [31]. With its thousands of years of history in treating diseases, TCM can be an important source of biologically active natural products. Chen reported that the first nonpeptidic GLP1R agonist Boc5 isolated from TCM had an antidiabetic efficacy in a diabetes mouse model (db/db mice)[32]. However, no further follow-up studies were published.

There are two major approaches in the field of natural product compounds research [33,34]. One is to isolate compounds and determine their chemical properties before subjecting them to

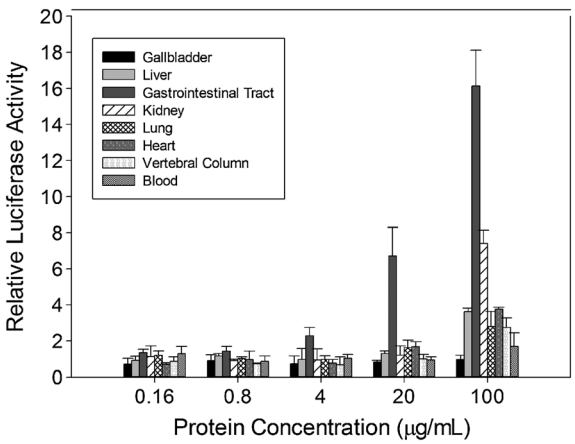

Figure 4. The glucagon-like peptide 1 receptor agonist activity of extractions from different tissues of *Ophisaurus harti*. The extracts were homogenized from organ tissues tested. The protein concentrations of extracts were measured using the bicinchoninic acid method. The gastrointestinal tract extract had the highest activity compared with the other organs in GLP1R$^+$ cell lines.

Journal of PeptideScience

Table 2. Summary of transcriptome analysis of *Ophisaurus harti*

Reads	
Total number	5 442 642
Total bases (bp)	49 537 780
Contigs	—
Total number	107 501
Mean length (bp)	605
Unigenes	—
Total number	58 959
Mean length (bp)	863

bp, basepairs.

biological activity studies; the other is to identify active compounds within a crude extract through purification steps guided by biological activity. This study used the second approach for it increased chance of identifying compounds that are either less abundant or with strong polarity, extreme hydrophobicity, large molecular weight, and so on. The original goal of this study was to identify less abundant small molecule GLP1R agonists, but a peptidic agonist was identified instead. The active protein GLP1 present in A089-147 would not have been identified using the first approach.

Initial efforts to further characterize A089-147 were hampered by the apparent heterogeneity of the active ingredients. Limited success was obtained in enriching the specific activity of A089-147 despite the different ion exchange and size exclusion chromatographies that had been attempted (data not shown). An alternative method for the identification of the active protein in A089-147 would use modern proteomics and the genomic method[35,36]. A comparison between the trypsin digested peptides from fractionated A089-147 and the *in silico* digested peptides derived from *O. harti's* complete gene sequence could be used to identify the specific gene coding for the target protein [26,37]. Unfortunately, this approach is currently not feasible because the *O. harti* genome has not been previously published. Over the last few years, the development of the next

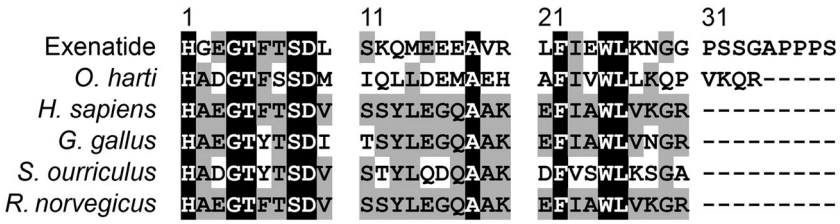

Figure 5. Sequence alignment of exenatide and glucagon-like peptide 1 (GLP1) peptides from different species. Amino acid sequences on black background represent six identical residues, and amino acid sequences on gray background represent three to five identical residues among the six species. GLP1 sequences are highly conserved among different species, in addition to exenatide and *Ophisaurus harti* GLP1, four out of 32 reported GLP1 sequences are included in Figure 5 [28].

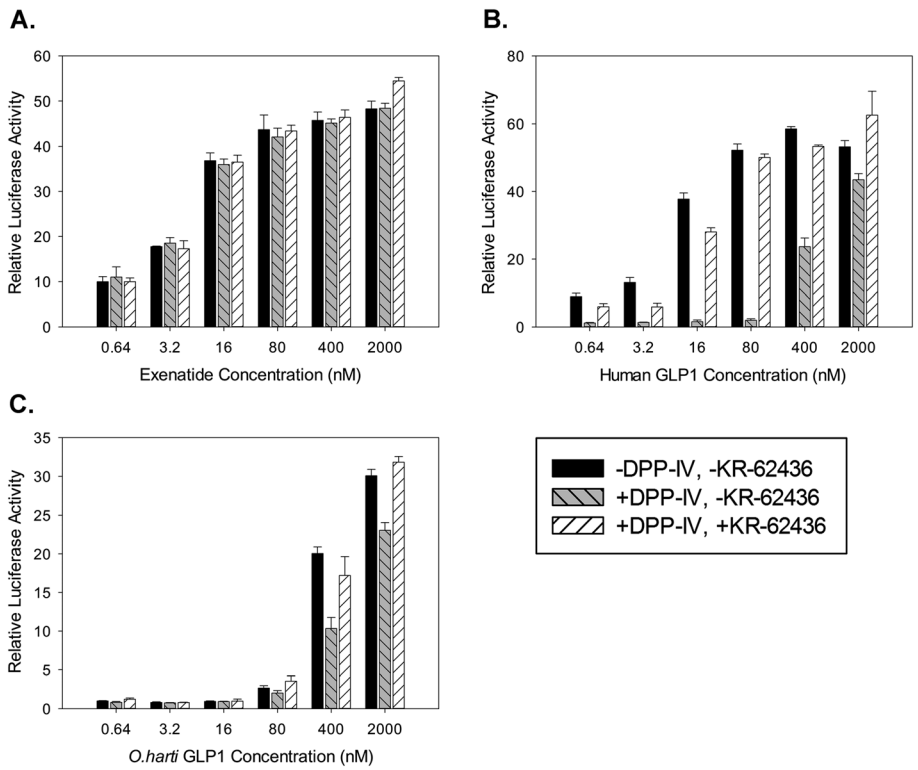

Figure 6. The dipeptidase IV (DPP-IV) resistance assays for (A) exenatide, (B) human glucagon-like peptide 1 (GLP1), (C) synthesized *Ophisaurus harti* GLP1. Peptides were pre-incubated with DDP-IV with or without DDP-IV inhibitor KR-62436 and then tested on GLP1R^{+} cells for remaining agonist activity. Human GLP1 and *Ophisaurus harti* GLP1 were inactivated by DPP-IV, although the presence of KR-62436 prevented this inactivation. Agonist activities from exenatide were DPP-IV resistant.

generation sequencing technologies has been used to produce *de novo* sequences from transcriptome templates [38]. Transcriptomes present a valuable resource for accelerating gene discovery by expanding gene families [39,40] and rapidly identifying transcripts involved in specific biological processes [41]. The transcriptome of *O. harti* was sequenced and annotated to positively identify the active ingredient of A089-147. The GLP1 sequence was identified and then confirmed by RT-PCR product sequencing.

The synthetic *O. harti* GLP1 activated the GLP1R$^+$ cells, and the potency of *O. harti* GLP1 is 10 to 20-fold lower than the human GLP1 [Figure 6, compare 6(C) with 6(B)]. In the reported Ala scanning study, side chains of human GLP1 at position His1, Gly4, Phe6, Thr7, and Asp9 were demonstrated to be important for GLP1 activity [42]. Substitution of Thr7 with a shorter side chain Ser residue in *O. harti* GLP1 may provide a less optimal distance for GLP interaction with GLP1R. Furthermore, in the human GLP1 crystal structure, four residues on the α-helix of human GLP1, Ala18, Ala19, Phe22, and Ile23 form part of the hydrophobic face, which provides hydrophobic interaction with the extracellular domain of GLP1R [43]. Substitution of Ala19 with Glu in *O. harti* GLP1 may decrease the hydrophobic interaction and result in decreased potency of *O. harti* GLP1.

Despite the novel *O. harti* GLP1 sequence identified in this study does not possess the DPP-IV resistance necessary to be clinically relevant, this study illustrated a comprehensive screening strategy to identify the novel GLP1R agonists from TCMs libraries. The novel *O. harti* GLP1 peptide yielded from this study confirmed broader application of TCMs libraries in active peptide/ingredient identification.

Acknowledgement

This work was supported by the Science and Technology Commission of Shanghai Municipality (STCSM) grants (No. 08PJ1404100 and 11DZ1921203) and the National Basic Research Program of China (973 Program, No. 2013CB530700).

References

1 Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004; **27**: 1047–53.
2 Yagi T, Nishi S, Hinata S, Murakami M, Yoshimi T. A population association study of four candidate genes (hexokinase II, glucagon-like peptide-1 receptor, fatty acid binding protein-2, and apolipoprotein C-II) with type 2 diabetes and impaired glucose tolerance in Japanese subjects. *Diabet. Med.* 1996; **13**: 902–7.
3 Fehmann HC, Jiang J, Pitt D, Schweinfurth J, Goke B. Ligand-induced regulation of glucagon-like peptide-I receptor function and expression in insulin-secreting beta cells. *Pancreas* 1996; **13**: 273–82.
4 Holz GG, Kuhtreiber WM, Habener JF. Pancreatic beta-cells are rendered glucose-competent by the insulinotropic hormone glucagon-like peptide-1(7–37). *Nature* 1993; **361**: 362–5.
5 Knop FK, Vilsboll T, Larsen S, Madsbad S, Holst JJ, Krarup T. No hypoglycemia after subcutaneous administration of glucagon-like peptide-1 in lean type 2 diabetic patients and in patients with diabetes secondary to chronic pancreatitis. *Diabetes Care* 2003; **26**: 2581–7.
6 Knudsen LB, Kiel D, Teng M, Behrens C, Bhumralkar D, Kodra JT, Holst JJ, Jeppesen CB, Johnson MD, de Jong JC, Jorgensen AS, Kercher T, Kostrowicki J, Madsen P, Olesen PH, Petersen JS, Poulsen F, Sidelmann UG, Sturis J, Truesdale L, May J, Lau J. Small-molecule agonists for the glucagon-like peptide 1 receptor. *Proc. Natl. Acad. Sci. U. S. A.* 2007; **104**: 937–42.
7 Widmann C, Dolci W, Thorens B. Agonist-induced internalization and recycling of the glucagon-like peptide-1 receptor in transfected fibroblasts and in insulinomas. *Biochem. J.* 1995; **310**: 203–14.
8 Deacon CF. Incretin-based treatment of type 2 diabetes: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors. *Diabetes Obes. Metab.* 2007; **9**(Suppl 1): 23–31.
9 Crepaldi G, Carruba M, Comaschi M, Del Prato S, Frajese G, Paolisso G. Dipeptidyl peptidase 4 (DPP-4) inhibitors and their role in type 2 diabetes management. *J. Endocrinol. Invest.* 2007; **30**: 610–4.
10 Eng J, Kleinman WA, Singh L, Singh G, Raufman JP. Isolation and characterization of exendin-4, an exendin-3 analogue, from Heloderma suspectum venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas. *J. Biol. Chem.* 1992; **267**: 7402–5.
11 Goke R, Fehmann HC, Linn T, Schmidt H, Krause M, Eng J, Goke B. Exendin-4 is a high potency agonist and truncated exendin-(9–39)-amide an antagonist at the glucagon-like peptide 1-(7–36)-amide receptor of insulin-secreting beta-cells. *J. Biol. Chem.* 1993; **268**: 19650–5.
12 Fineman MS, Shen LZ, Taylor K, Kim DD, Baron AD. Effectiveness of progressive dose-escalation of exenatide (exendin-4) in reducing dose-limiting side effects in subjects with type 2 diabetes. *Diabetes Metab. Res. Rev.* 2004; **20**: 411–7.
13 Kolterman OG, Buse JB, Fineman MS, Gaines E, Heintz S, Bicsak TA, Taylor K, Kim D, Aisporna M, Wang Y, Baron AD. Synthetic exendin-4 (exenatide) significantly reduces postprandial and fasting plasma glucose in subjects with type 2 diabetes. *J. Clin. Endocrinol. Metab.* 2003; **88**: 3082–9.
14 Egan JM, Clocquet AR, Elahi D. The insulinotropic effect of acute exendin-4 administered to humans: comparison of nondiabetic state to type 2 diabetes. *J. Clin. Endocrinol. Metab.* 2002; **87**: 1282–90.
15 Maoshing N. The Yellow Emperor's Classic of Medicine: A New Translation of the Neijing Suwen with Commentary. Boston, MA: Shambhala; 1995.
16 Liang XC, Guo SS, Qian ZF, Hong G, Tao LH, Wang Y, Zhu SY. Effects of Chinese medicinal herbs on hemorheology in diabetics. *Proc. Chin. Acad. Med. Sci. Peking Union Med. Coll.* 1989; **4**: 135–8.
17 Cheng JT, Liu IM, Chi TC, Su HC, Chang CG. Metformin-like effects of Quei Fu Di Huang Wan, a Chinese herbal mixture, on streptozotocin-induced diabetic rat. *Horm. Metab. Res.* 2001; **33**: 727–32.
18 Vray M, Attali JR. Randomized study of glibenclamide versus traditional Chinese treatment in type 2 diabetic patients. Chinese-French Scientific Committee for the Study of Diabetes. *Diabete Metab.* 1995; **21**: 433–9.
19 Zheng X, Hu Y, Liu J, Ouyang K. Screening of active compounds as neuromedin U2 receptor agonist from natural products. *Bioorg. Med. Chem. Lett.* 2005; **15**: 4531–5.
20 Fitzgerald LR, Mannan IJ, Dytko GM, Wu HL, Nambi P. Measurement of responses from Gi-, Gs-, or Gq-coupled receptors by a multiple response element/cAMP response element-directed reporter assay. *Anal. Biochem.* 1999; **275**: 54–61.
21 Hu JF, Garo E, Yoo HD, Cremin PA, Zeng L, Goering MG, O'Neil-Johnson M, Eldridge GR. Application of capillary-scale NMR for the structure determination of phytochemicals. *Phytochem. Anal.* 2005; **16**: 127–33.
22 Hu JF, Yoo HD, Williams CT, Garo E, Cremin PA, Zeng L, Vervoort HC, Lee CM, Hart SM, Goering MG, O'Neil-Johnson M, Eldridge GR. Miniaturization of the structure elucidation of novel natural products - two trace antibacterial acylated caprylic alcohol glycosides from Arctostaphylos pumila. *Planta Med.* 2005; **71**: 176–80.
23 Kim KR, Rhee SD, Kim HY, Jung WH, Yang SD, Kim SS, Ahn JH, Cheon HG. KR-62436, 6-{2-[2-(5-cyano-4,5-dihydropyrazol-1-yl)-2-oxoethylamino] ethylamino}nicotinonitrile, is a novel dipeptidyl peptidase-IV (DPP-IV) inhibitor with anti-hyperglycemic activity. *Eur. J. Pharmacol.* 2005; **518**: 63–70.
24 Mojsov S, Weir GC, Habener JF. Insulinotropin: glucagon-like peptide I (7–37) co-encoded in the glucagon gene is a potent stimulator of insulin release in the perfused rat pancreas. *J. Clin. Invest.* 1987; **79**: 616–9.
25 Kreymann B, Williams G, Ghatei MA, Bloom SR. Glucagon-like peptide-1 7–36: a physiological incretin in man. *Lancet* 1987; **2**: 1300–4.
26 Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 2009; **10**: 57–63.
27 Wei W, Qi X, Wang L, Zhang Y, Hua W, Li D, Lv H, Zhang X. Characterization of the sesame (Sesamum indicum L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics* 2011; **12**: 451.
28 Kieffer TJ, Habener JF. The glucagon-like peptides. *Endocr. Rev.* 1999; **20**: 876–913.
29 Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SF, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R,

Kochva E. Early evolution of the venom system in lizards and snakes. *Nature* 2006; **439**: 584–8.
30 Butler MS. The role of natural product chemistry in drug discovery. *J. Nat. Prod.* 2004; **67**: 2141–53.
31 Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 2007; **70**: 461–77.
32 Chen D, Liao J, Li N, Zhou C, Liu Q, Wang G, Zhang R, Zhang S, Lin L, Chen K, Xie X, Nan F, Young AA, Wang MW. A nonpeptidic agonist of glucagon-like peptide 1 receptors with efficacy in diabetic db/db mice. *Proc. Natl. Acad. Sci. U. S. A.* 2007; **104**: 943–8.
33 Crawford AD, Liekens S, Kamuhabwa AR, Maes J, Munck S, Busson R, Rozenski J, Esguerra CV, de Witte PA. Zebrafish bioassay-guided natural product discovery: isolation of angiogenesis inhibitors from East African medicinal plants. *PLoS One* 2011; **6**: e14694.
34 Bugni TS, Richards B, Bhoite L, Cimbora D, Harper MK, Ireland CM. Marine natural product libraries for high-throughput screening and rapid drug discovery. *J. Nat. Prod.* 2008; **71**: 1095–8.
35 Service RF. Proteomics. High-speed biologists search for gold in proteins. *Science* 2001; **294**: 2074–7.
36 Smaglik P. Genomics initiative to decipher 10,000 protein structures. *Nature* 2000; **407**: 549.
37 Graham IA, Besser K, Blumer S, Branigan CA, Czechowski T, Elias L, Guterman I, Harvey D, Isaac PG, Khan AM, Larson TR, Li Y, Pawson T, Penfield T, Rae AM, Rathbone DA, Reid S, Ross J, Smallwood MF, Segura V, Townsend T, Vyas D, Winzer T, Bowles D. The genetic map of Artemisia annua L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science* 2011; **327**: 328–31.
38 Riggins CW, Peng Y, Stewart CN, Jr., Tranel PJ. Characterization of de novo transcriptome for waterhemp (Amaranthus tuberculatus) using GS-FLX 454 pyrosequencing and its application for studies of herbicide target-site genes. *Pest Manag. Sci.* 2011; **66**: 1042–52.
39 Bourdon V, Naef F, Rao PH, Reuter V, Mok SC, Bosl GJ, Koul S, Murty VV, Kucherlapati RS, Chaganti RS. Genomic and expression analysis of the 12p11-p12 amplicon using EST arrays identifies two novel amplified and overexpressed genes. *Cancer Res.* 2002; **62**: 6218–23.
40 Cheung F, Win J, Lang JM, Hamilton J, Vuong H, Leach JE, Kamoun S, Andre Levesque C, Tisserat N, Buell CR. Analysis of the Pythium ultimum transcriptome using Sanger and Pyrosequencing approaches. *BMC Genomics* 2008; **9**: 542.
41 Rismani-Yazdi H, Haznedaroglu BZ, Bibby K, Peccia J. Transcriptome sequencing and annotation of the microalgae Dunaliella tertiolecta: pathway description and gene discovery for production of next-generation biofuels. *BMC Genomics* 2011; **12**: 148.
42 Adelhorst K, Hedegaard BB, Knudsen LB, Kirk O. Structure-activity studies of glucagon-like peptide-1. *J. Biol. Chem.* 1994; **269**: 6275–8.
43 Underwood CR, Garibay P, Knudsen LB, Hastrup S, Peters GH, Rudolph R, Reedtz-Runge S. Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J. Biol. Chem.* 2010; **285**: 723–30.

