Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
ACE2 orthologues in non-mammalian vertebrates 
(Danio, Gallus, Fugu, Tetraodon and Xenopus)

Chih-Fong Chou*, Chay Boon Loh, Yik Khoon Foo, Shuo Shen, Burtram C. Fielding 1, Timothy H.P. Tan, Sehaam Khan 2, Yue Wang, Seng Gee Lim, Wanjin Hong, Yee-Joo Tan, Jianlin Fu

Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673

Received 27 September 2005; received in revised form 7 March 2006; accepted 7 March 2006

Available online 5 April 2006

Received by O. Clay

Abstract

Angiotensin-converting enzyme 2 (ACE2), a newly identified member in the renin–angiotensin system (RAS), acts as a negative regulator of ACE. It is mainly expressed in cardiac blood vessels and the tubular epithelia of kidneys and abnormal expression has been implicated in diabetes, hypertension and heart failure. The mechanism and physiological function of this zinc metallopeptidase in mammals are not yet fully understood. Non-mammalian vertebrate models offer attractive and simple alternatives that could facilitate the exploration of ACE2 function. In this paper we report the in silico analysis of Ace2 genes from the Gallus (chicken), Xenopus (frog), Fugu and Tetraodon (pufferfish) genome assembly databases, and from the Danio (zebrafish) cDNA library. Exon ambiguities of Danio and Xenopus Ace2 were solved by RT-PCR and 3′RACE. Analyses of the exon–intron structures, alignment, phylogeny and hydrophilicity plots, together with the conserved synteny among these vertebrates, support the orthologous relationship between mammalian and non-mammalian ACE2s. The putative promoters of Ace2 from human, Tetraodon and Xenopus tropicalis drove the expression of enhanced green fluorescent protein (EGFP) specifically in the heart tissue of transgenic Xenopus thus making it a suitable model for future functional genomic studies. Additionally, the search for conserved cis-elements resulted in the discovery of WGATAR motifs in all the putative Ace2 promoters from 7 different animals, suggesting a possible role of GATA family transcriptional factors in regulating the expression of Ace2.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Angiotensin; Converting; Transgenic; Heart; GATA

1. Introduction

The renin–angiotensin system (RAS) plays a key role in the regulation of cardiovascular and renal functions. Renin stimulates
contains two HEXXH+E zinc binding motifs whereas ACE2 has only one for its monopeptidase activity. In contrast to the widely distributed ACE in tissues, the expression of ACE2 is more restricted, being found mainly on the endothelial cells of the arteries, arterioles and venules in kidney and heart (Donoughue et al., 2000; Tipnis et al., 2000; Oudit et al., 2003). ACE2 is also expressed in the renal tubular epithelium, the vascular smooth muscle cells of the intrarenal arteries, coronary blood vessels (Donoughue et al., 2000) and adult Leydig cells of the testis (Douglas et al., 2004).

From studies of kidney and heart disease caused by unregulated expression of ACE2, it was suggested that ACE2 has a renoprotective role in diabetes as both Ace2 mRNA and protein levels were decreased by ∼ 50% in diabetic renal tubules. This reduction could be prevented by ACE inhibitor therapy (Burrell et al., 2004). Severe impairment in myocardial contractility, with accompanying higher AngII levels, has been observed in ACE2 null mice, and this cardiac defect was rescued completely by genetic ablation (double knockouts) of ACE (Crackower et al., 2002; Oudit et al., 2003). The antagonistic relationship between ACE and ACE2 modulates the balance between AngII (vasopressor) and Ang 1–7 (vasodilator), which plays a significant role in regulating renal and cardiovascular functions. ACE2 has recently been identified as the functional receptor of severe acute respiratory syndrome (SARS) coronavirus (CoV) (Li et al., 2003) and its expression was located on the surface of lung alveolar epithelial cells and enterocytes of the small intestine (Hamming et al., 2004). SARS-CoV infections and the spike (S) protein reduce the expression of ACE2, and injection of the S protein into mice exacerbates acute lung failure in vivo (Kuba et al., 2005). Furthermore, ACE2 and the AngII type 2 receptor protect mice from severe acute lung injury whereas other components of the RAS (including ACE, AngII, and AngII type 1a receptor) promote disease pathogenesis (Imai et al., 2005). These findings provide a possible therapeutic role for ACE2 in acute lung injury affecting millions of people worldwide every year. As the receptor of SARS-CoV, ACE2 was shown to mediate the binding between a Vero E6–ACE2 expressing cell line and the recombinant S protein expressed on the surface of CHO cells, even under high stringency washing conditions (Chou et al., 2005).

Overexpression of ACE2 in the hearts of transgenic mice results in lethal ventricular arrhythmia and subsequent heart failure. This has been shown to be associated with the disruption of gap junction formation (Donoghue et al., 2003). The high occurrence of sudden death of transgenic mice correlates with the expression levels of ACE2. For ACE2 to play the appropriate physiological role in RAS and in the maintenance of properly functioning heart or kidney, the expression requires not only the correct level but also the tissue specificity. According to the World Health Organization predictions cardiovascular diseases will be the leading cause of death by the year 2020, thus it is imperative to include ACE2 as a potential target for developing therapies to control such diseases. Currently, mouse and rat models are available for the study of ACE2, however non-mammalian vertebrates such as frog, zebrafish (Lohr and Yost, 2000) and chicken are particularly useful in the studies for dissecting heart development at early stages. Frog heart contains two atria and a single ventricle, and is a good tool for studying congenital heart diseases. In this paper, we describe a comparison between ACE2 sequences of chicken, frog, pufferfish and zebrafish. We also show the expression of EGFP in transgenic Xenopus hearts, as driven by Ace2 promoters from human, Tetraodon and Xenopus.

2. Materials and methods

2.1 Identification of non-mammalian vertebrates Ace2 genes

Human ACE2 protein sequence (NCBI accession no. AF291820) was used to BLAST the Fugu database http://fugubiology.qmul.ac.uk/blast using the “blastn” algorithm and partially identified Fugu Ace2 coding sequence was used to BLAST search at http://www.ncbi.nlm.nih.gov/BLAST for Tetraodon Ace2. Full length Ace2 genes of Fugu and Tetraodon were completed by using the program GENSCAN (http://genes.mit.edu/GENSCAN.html) to locate coding sequences from the genomic DNA sequences. The incomplete Danio rerio Ace2 mRNA (NCBI accession no. BC085667) was completed by sequencing the RT-PCR fragments of the 13th and 18th exons. The partial Gallus gallus and Xenopus tropicalis Ace2 gene sequences were found by BLAST searches against the database of University of California, Santa Cruz Genome Bioinformatics at http://genome.ucsc.edu using the human protein sequence. The chicken was completed in the same way as the pufferfish genes, however, exon 18 of Xenopus Ace2 was confirmed by 3’RACE.

2.2 Danio and Xenopus mRNA preparation, RT-PCR and 3’ RACE

The Zebrafish Ace2 13th exon was RT-PCR amplified from RNA mix of 4 to 10-hpf (hours post fertilization) Zebrafish using Qiagen One-Step RT-PCR Kit (Valencia, CA). Primers 5′-CATGAGTTGGTGAAGGGG-3′ and 5′-ATGCTATTGCGTTCCAGGTT-3′ were used and the RT-PCR was performed according to manufacturer’s instructions. Fragments of 156 bp were extracted from the gel (Qiagen Qiaquick Gel Extraction kit), and cloned using Invitrogen (Carlsbad, CA) TOPO TA Cloning Kit and then sequenced. The PCR primer pair: 5′-TAGAAGCTTGGCGAGAACATCTTAAATTG-3′ and 5′-TAGGGTACCTTACAGTCCTGTTTGTTCAATG-3′ was used to clone the 409 bp fragment (containing the 18th exon of Danio Ace2) in a similar manner.

An adult male X. tropicalis was sacrificed and its heart was surgically removed for RNA extraction using Trizol reagent (Invitrogen) according to the manufacturer’s protocol. The purified total RNA was reverse transcribed into cDNA using Superscript RT (Gibco BRL, Grand Island, NY) and anchor primer, 5′-TTGACCACCGTGATCATGTCGACCTTAAAAATTG-3′ and 5′-TAGGTACCTTACAGTCCTGTTTGTTCAATG-3′ was used to clone the 409 bp fragment (containing the 18th exon of Danio Ace2) in a similar manner.

C.-F. Chou et al. / Gene 377 (2006) 46–55 47
and the reverse primer 5′-CCTTGACCACGCGTATC-GATGCTGTTTT-3′ were used to amplify the 3′ RACE fragment. TOPO cloning was employed to clone fragments around 1.1 kb for sequencing.

2.3. Sequencing analysis, alignment, phylogenetics and hydrophilicity plot

Human and mouse Ace2 sequences (protein [mouse ACE2 NCBI accession no.AAH26801], cDNA [human NCBI accession no. BC039902 and mouse NCBI accession no. BC026801] and genomic DNA [Supplementary data 1]) were retrieved from the NCBI database. Genomic DNA and cDNA of Ace2 from human, mouse, chicken, frog and pufferfish were analyzed for comparing the exon–intron organization. Alignments of the Ace2 protein sequences of all animals were generated using the program ClustalW. The phylogenetic analysis of ACE2s and ACEs were obtained using the program MegAlign (DNASTar, Madison, WI) with the default ClustalW parameters. The accession nos. of ACEs are CAG04404 (Tetraodon), Q10751 (chicken), AAR03504 (human) and P09470 (mouse). The hydrophilicity plots of the ACE2 sequences were analyzed using DNASTar Protean with the default Kyte and Doolittle hydrophilicity parameters.

2.4. Construction of Ace2 promoter–EGFP constructs for transgenic Xenopus

The intergenic region between Tetraodon Ace2 and Nhs as well as the 5′UTR of human (0.6 kb) and Xenopus (0.5 kb) were used as putative Ace2 promoters. They were amplified by PCR using Tetraodon genomic DNA (isolated from green spotted puffer), human genomic DNA (isolated from ATCC CCL-171 MRC-5 cell line) and Xenopus genomic DNA (isolated from X. tropicalis) as templates. The following PCR primer pairs: 5′-GCGGATCCCTTACTTCCTGAGCCGAGG-3′ and 5′-GCGGATCCCGTGCACGCCAGAAATAACG-3′ were for the amplification of the 2.8 kb intergenic region of Tetraodon, 5′-GCGGTACCCAGACCGAGACTCAGTCTC-3′ plus 5′-GCGGATCCGGTCCCTCCTGAGCAGCCGATC-3′ were for the amplification of human 0.6 kb putative promoter and 5′-GCGGATCCAGGTGAAACACATCGTGACG-3′ with 5′-GCGGATCCCTTACTTCCTGAGCCGAGG-3′ were for Xenopus 0.5 kb putative promoter. All fragments were TOPO-cloned and checked by sequencing, and then inserted into pEGFP-1 (Clontech, Palo Alto, CA) at the BamHI site for the Tetraodon fragment or KpnI and BamHI sites for human and Xenopus putative promoters. XhoI and BglII restriction enzymes were used to obtain the DNA fragments from all constructs, each fragment containing the Ace2 putative promoter with the reporter EGFP and the polyadenylation signal. Fragments purified from the gel by Qiagen Qiaquick Gel Extraction kit were microinjected into Xenopus laevis eggs.

2.5. Microinjection

A modified restriction enzyme-mediated integration was used to generate transgenic X. laevis tadpoles (Kroll and Amaya, 1996; Sparrow et al., 2000). Briefly, unfertilized eggs were squeezed from female frog primed with HCG (human chorionic gonadotropin; Sigma-Aldrich, St. Louis, MO) and dejellied by 2% L-cysteine (W/V) in 1× MMR (100 mmol/L NaCl, 2 mmol/L KCl, 1 mmol/L MgCl2, 2 mmol/L CaCl2 and 5 mmol/L HEPES at pH 7.5). Dejellied eggs were transferred to 60 mm Petri-dish coated with 2% solidified agarose and filled with 0.4× MMR +6% Ficoll 400. Meanwhile, 2.5 μL of sperm nuclei (stock concentration of 1×10⁷/μL) and 2.5 μL of linearized DNA (use 100 ng of DNA for every 3 kb) were mixed and left at room temperature for 10 min. The sperm nuclei and DNA mix were then diluted with 495 μL Sperm Dilution Buffer. Injections were carried out with a 60 μL ID needle and at a constant pump rate of 0.7 μL/min (Harvard Apparatus DHD 2000 Infusion). Injected eggs were incubated at 16 °C for 3 h. Normal developed 4-cells stage embryos were selected and cultured in 0.1× MMR +6% Ficoll 400 +50 μg/mL gentamycin at 16 °C. These embryos are cultured until the 15–20th stage before transferring to 0.1× MMR buffer containing 50 μg/mL gentamycin and left at room temp. Successfully transgene tadpoles can be identified by EGFP production under a fluorescent microscope and recorded by digital camera.

2.6. Identification of potential transcription factor binding sites in Ace2 gene promoter

One kilobase 5′UTR sequences upstream of the ATG codon of 7 vertebrate Ace2 genes (human, mouse, chicken, Xenopus, Fugu, Tetraodon and zebrafish) were searched athttp://www.cbrc.jp/research/db/TFSEARCH for transcription factor binding sites.

3. Results

3.1. Non-mammalian Ace2 genes

Fugu scaffold 2251 (GenBank accession no. CAAB 01002251.1) containing Fugu rubripes Ace2 was BLAST identified using the human ACE2 protein sequence. To confirm the finding, we checked human Ace2 neighbor genes for Fugu homologs which were located in the same scaffold using tblastn. A Nance–Horan Syndrome (Nhs) gene (present at the human Ace2 locus) homologue was also found in Fugu scaffold 2251. The gene was connected to Fugu Ace2 in a head-to-head orientation without any genes in the intergenic region. Ace2 and Nhs genes of Tetraodon nigroviridis were found in chromosome 5 (NCBI accession no. CAAE01014581) in the same orientation as Fugu genes. Reference pufferfish ACE2 sequences were used to assist in the identification of homolog genes from chicken, frog and zebrafish. The full length ACE2 sequences of Fugu (803 amino acids, see Supplementary data 1), Tetraodon (803 a.a.; Supplementary data 1) and Gallus (808 a.a.; Supplementary data 1) were identified in silico from available databases. Since teleosts are known to contain 2 copies of several single-copy mammalian genes due to a whole-genome duplication early during the evolution of teleosts (Christoffels et al., 2004), we searched the Fugu and Tetraodon genomes using the identified ACE2 protein sequences to determine whether there is a second copy. However, our search indicated that
ACE2 is a single copy gene in these teleosts, as are some other genes such as the Fugu erythropoietin gene (Chou et al., 2004). Incomplete Danio ACE2 protein sequence was obtained using BLAST (blastp) at NCBI with the Fugu ACE2 protein sequence. It was described as a 785 amino acid protein with unknown function (NCBI accession no. AAH85667). When

Table 1
Exon–intron structures

| Exons (coding sequences) | Introns (intronic regions) |
|-------------------------|---------------------------|
| hAce2 and mAce2         | hAce2                     |
| gAce2                   | mAce2                     |
| xAce2                   | gAce2                     |
| fAce2 and tAce2         | xAce2                     |
| fAce2                   | fAce2                     |
| tAce2                   | tAce2                     |

| 1 | 186 | 183 | 189 | 195 | 5722 | 12,203 | 2220 | 21,076 | 86 | 86 |
| 2 | 159 | 159 | 159 | 159 | 2522 | 3297 | 1785 | 679 | 100 | 91 |
| 3 | 94  | 94  | 94  | 94  | 372  | 1257 | 571  | 1901 | 335 | 315 |
| 4 | 144 | 144 | 144 | 144 | 2256 | 2762 | 456  | 939  | 132 | 150 |
| 5 | 113 | 113 | 113 | 116 | 1485 | 1402 | 681  | 628  | 75  | 79 |
| 6 | 106 | 109 | 109 | 106 | 2180 | 1538 | 708  | 292  | 353 | 511 |
| 7 | 98  | 98  | 89  | 89  | 4084 | 3782 | 1428 | 583  | 102 | 88 |
| 8 | 170 | 170 | 170 | 173 | 2905 | 1609 | 867  | 531  | 611 | 601 |
| 9 | 227 | 227 | 227 | 227 | 2278 | 2558 | 604  | 1037 | 348 | 260 |
| 10| 145 | 145 | 145 | 145 | 2200 | 1561 | 1604 | 1312 | 126 | 103 |
| 11| 99  | 99  | 99  | 99  | 1043 | 1088 | 668  | 169  | 152 | 95 |
| 12| 123 | 123 | 123 | 123 | 404  | 428  | 319  | 493  | 148 | 86 |
| 13| 173 | 173 | 179 | 179 | 1270 | 1492 | 704  | 820  | 796 | 867 |
| 14| 59  | 59  | 62  | 59  | 2468 | 2887 | 3951 | 4212 | 127 | 129 |
| 15| 101 | 101 | 101 | 98  | 1356 | 2065 | 951  | 1385 | 88  | 95 |
| 16| 117 | 117 | 111 | 117 | 2034 | 2925 | 1197 | 290  | 126 | 95 |
| 17| 195 | 195 | 195 | 195 | 2010 | 1665 | 609  | 1210 | 151 | 113 |
| 18| 106 | 115 | 268 | 91  | 36,589 | 44,519 | 19,323 | 37,557 | 3856 | 3764 |

The number of nucleotides in each exon and intron are shown.

a Ace2s are prefixed by h (human), m (mouse), g (chicken), x (Xenopus), f (Fugu) and t (Tetraodon).

b Highlighted nucleotide numbers indicating the exon among all Ace2s are same.

c Total base pairs.

Fig. 1. Ace2–Nhs loci of human (chromosome [chr] X), chicken (chr 1), Xenopus (scaffold 403, 38 and 56), Fugu (scaffold 2251) and Tetraodon (chr 5). Genes and their orientations are shown as arrows. The dashed line in Xenopus locus indicates the discontinuity and distinguishes the three scaffolds. The gene orders were obtained from the UCSC Genome Browsers (http://genome.ucsc.edu). Locus positions in the chromosome or scaffold are underlined.
this sequence was aligned with human and pufferfish ACE2s, it was found to contain 11 extra a.a. residues at the end of the 13th exon. Also, exon 18 was absent from the sequence. Primers were designed according to the *Danio* cDNA sequence (NCBI accession no. BC085667) for RT-PCR, and PCR products from exons 13 and 18 (based on the alignment) were sequenced. The 33 nucleotides (coding for the 11 a.a. peptide) in the BC085667 sequence and a G at the end of the 17th exon were absent in our sequence data. The extra G causes a frame shift that results in a premature stop signal, 3 codons downstream of exon 17. The full-length *Danio* ACE2 sequence was then identified as an 807 a.a. (Supplementary Data 1) polypeptide. The partial sequence of frog ACE2 was found in the *X. tropicalis* scaffold 403 by BLAST, and GENESCAN was used to identify the 18th

![Fig. 2. The hydrophilicity plots for ACE2s from human, mouse, chicken, *Xenopus*, *Fugu*, *Tetraodon* and zebrafish generated using Kyte and Doolittle hydrophilicity parameters. The location of the N-terminal amino acid (a.a.) residue of the mature ACE2s is indicated by arrows. □ denotes the transmembrane domains.](image-url)

Fig. 2. The phylogenetic analysis of the ACE2 protein sequences, obtained using the program MegAlign with the default ClustalW parameters. The source of ACE2s and ACEs are prefixed by h (human), m (mouse), g (chicken), x (*Xenopus*), f (*Fugu*), t (*Tetraodon*) and z (zebrafish).
exon which contains 162 additional nucleotides compared to the 18th exon of human Ace2 (Table 1). To confirm the sequence, we designed 3' RACE primers based on exon 17. We found that the polyadenylation signal was 680 bp downstream of the stop codon, and the frog Ace2 encodes 859 a.a. (Supplementary Data 1) residues with an extra 54 a.a (compared to human ACE2) located in the cytoplasmic domain.

3.2. Exon–intron structure of Ace2 genes

We determined the exon–intron structure of the chicken, frog and pufferfish Ace2 genes by mapping the cDNAs (Supplementary data 1) to their respective genomic sequence. The consensus GT and AG splice donor and acceptor sequences used for the mappings are suitable for all the genes in Table 1. All of the vertebrate Ace2 genes contain 18 exons and 17 introns, similar to the mammalian Ace2s. The lengths of all corresponding exons of Ace2 are identical between human and mouse and between Fugu and Tetraodon. Among all the Ace2s shown in Table 1, exons 2, 3, 4, 9, 10, 11, 12 and 17 are conserved in length, whereas the lengths of exons 1 and 18 are different in all Ace2s. The Xenopus has the longest Ace2 with 2477 bp (Supplementary data 1) and the shortest 2409 bp are found in the pufferfish (Supplementary data 1). Pufferfishes are known for their compact genomes, their genes are densely found in the pufferfish (Supplementary data 1). Pufferfishes are represented by important animal models studied today that have evolved from the Osteichthyes (Bony fishes) lineage. We used ACE2 protein sequences from these 7 animal models together with ACE sequences of the phylogram supported the standard classification of the ACE2 sequences of all mammalian and non-mammalian Ace2 genes. Among all 7 Ace2s, Fugu and Tetraodon share the highest sequence identity of 88% while the human and mouse proteins are slightly lower (82%). The identities between human and non-mammalian ACE2s are from 56% (pufferfish) to 66% (chicken), between pufferfish and the other 5 ACE2s are around 53% to 57% (frog, human, mouse and chicken) and 65% (zebrafish). Although the identities among ACE2s varies from 53% to 88%, the hydrophilicity plots of all 7 ACE2s are very similar (Fig. 2), with the exception of the Xenopus ACE2, which has a larger cytoplasmic domain. Highly hydrophobic a.a. residues at the N-termini make up the signal peptides of ACE2s and the N-terminal a.a. residue of the mature proteins is indicated. The transmembrane domains are located at the hydrophobic region around a.a. residue 750, and the rest of the downstream sequence forms the small cytoplasmic domain.

3.3. Conserved synteny between Ace2 genes from different vertebrates

We found that Nhs and Ace2 genes are arranged with head-to-head juxtaposition in both Fugu and Tetraodon genomic sequences. The intergenic regions between them in the Fugu and Tetraodon genomes are 2.9 and 2.8 kb respectively. This Ace2–Nhs locus was used to study the synteny relationship among the loci in human, chicken, frog and pufferfishes. Excluding the Ace2 and Nhs genes, there are 9 other genes in the human locus (see http://genome.ucsc.edu and also for the identification of other loci) within a 1775 kb fragment in chromosome X (Fig. 1). The chicken locus contains 8 human orthologs (Tmem27, U2af1l2, Apis2, Grpr, Ctps2, Syap1, Rbbp7 and Reps2) except the gene Ca5b, and the fragment spans 903 kb in chromosome 1. The locus of X. tropicalis is separated in 3 parts according to the available information from the genome assembly program. We found 3 scaffolds (403, 38 and 56) that contain the 8 chicken orthologs as shown in Fig. 1. There are other genes downstream of Ctps2 in scaffold 403, and genes in both flanking regions of Syap1 in scaffold 38 as well as upstream of Rbbp7 in scaffold 56. The locus is either discontinuous in the Xenopus genome or forms a much bigger fragment. The Tetraodon Ace2 locus is in chromosome 5 while the location of the Fugu locus is currently unknown.

3.4. Comparisons of ACE2 sequences

Supplementary Data 2 presents the alignment of ACE2s from human, mouse, chicken, frog and teleosts (Fugu, Tetraodon and zebrafish). ACE2 is a zinc metallo-carboxypeptidase and it is distinct from ACE as it carries only a single catalytic domain to perform the monopeptidase activity. As seen from the alignment, the HEXXH+E zinc binding consensus motif was found to be conserved in exon 9 of all mammalian and non-mammalian Ace2 genes. Among all 7 ACE2s, Fugu and Tetraodon share the highest identity of 88% while the human and mouse proteins are slightly lower (82%). The identities between human and non-mammalian ACE2s are from 56% (pufferfish) to 66% (chicken), between pufferfish and the other 5 ACE2s are around 53% to 57% (frog, human, mouse and chicken) and 65% (zebrafish). Although the identities among ACE2s varies from 53% to 88%, the hydrophilicity plots of all 7 ACE2s are very similar (Fig. 2), with the exception of the Xenopus ACE2, which has a larger cytoplasmic domain. Highly hydrophobic a.a. residues at the N-termini make up the signal peptides of ACE2s and the N-terminal a.a. residue of the mature proteins is indicated. The transmembrane domains are located at the hydrophobic region around a.a. residue 750, and the rest of the downstream sequence forms the small cytoplasmic domain.

3.5. Ace2 putative promoters are tissue specifically active in the heart of transgenic frogs

ACE2 expression in the testis, kidney and heart had been demonstrated using Northern blot analysis (Donoghue et al., 2000). To analyze tissue specificity, we used the putative promoter of Tetraodon and human Ace2 to drive the expression of EGFP in transgenic frogs. In addition, to confirm the Ace2 of X. tropicalis (regarding the larger cytoplasmic domain) we checked the promoter tissue specificity as well. All three constructs were linearized and mixed with X. laevis sperm nuclei before microinjection into the eggs. After 5 days, around 10% of transgenic tadpoles expressed EGFP on the surface of heart ventricle or truncus arteriosus when visualized with a fluorescent microscope (Fig. 4 A and B). From our observation, for all constructs, the heart of transgenic frogs is the only tissue expressing EGFP. However, the nonspecific fluorescence from the gastrointestinal tracts (Fig. 4A) might mask expression in other tissues. The nonspecific fluorescence is also present in
wild type frogs. A movie of a beating fluorescent heart from a transgenic frog (using human 0.6 kb putative Ace2 promoter to drive the expression of EGFP) is presented as Supplementary data 3.

3.6. In silico prediction of the cis-acting motif for transcriptional factor GATA

Since the promoters of human and pufferfish Ace2 genes were found to drive expression in the heart of transgenic Xenopus (Fig. 4), we then compared the promoter sequences of Ace2 genes from 7 vertebrates to search for transcriptional factor binding sites that might be responsible for this tissue specific expression. The DNA sequences of 7 putative Ace2 promoters (1 kb upstream of ATG codon) were analyzed at http://www.cbrc.jp/research/db/TFSEARCH, and found to contain a conserved cis-element WGATAR which may serve as GATA transcription factor binding site (Ikonomi et al., 2000). The conserved motif WGATAR (T/AGATAA/G or reverse complement sequences C/TTATCT/A) was found between nt −640 and −230 (where the A of the ATG codon is +1) in all the 7 Ace2 promoters (Fig. 5). The total number of this specific motif was found to vary between species, 1 in human, mouse and Xenopus, 2 in chicken, 3 in zebrafish, 5 in Fugu and 6 in Tetraodon. These conserved motifs may provide possible targets for one or more of the GATA family transcriptional factors.

4. Discussion

In humans, it is predicted that cardiovascular diseases will be the most common cause of death worldwide by 2020. Our body uses the RAS to regulate systemic blood pressure, thereby maintaining stable total blood circulation. However, chronic high blood pressure results in hypertension which is the main cause of cardiovascular diseases. Since the discovery of ACE involvement in the production of the vasopressor AngII, numerous studies have focused on finding either ACE inhibitors or AngII receptor blockers as a means to counteract hypertension. ACE2, an antagonist of ACE in RAS, inactivates AngII to reduce the blood pressure, especially in heart arteries. To augment our understanding of the overall function of ACE2, we compared the structure of Ace2 genes from various non-mammalian vertebrate genomes. We searched the Fugu (Aparicio et al., 2002) and Tetraodon (Jaillon et al., 2004) genome assembly databases because they are well documented. Together with Ace2s from pufferfishes, we also identified the gene from other selected vertebrates (chicken, frog and zebrafish) that are important in evolutionary and developmental studies.

Fig. 4. Ace2 promoters drive the tissue specific expression of EGFP on the heart surface of transgenic Xenopus. (A) The expression was driven by 0.5 kb Xenopus Ace2 promoter region (lower panels) and 0.6 kb human Ace2 promoter region (upper panels). EGFP was visualized using a fluorescent microscope. The left panels were photos taken under brightfield and the right panels were taken under settings for EGFP fluorescence. The ventricle is shown by arrows and the truncus arteriosus by dashed arrows. (B) Fluorescenced hearts of transgenic Xenopus using the putative promoter of Tetraodon are shown and indicated by arrows. Photos were taken under filtered UV and dimmed light for presenting the fluorescenced heart in entire tadpoles.
The sizes of Ace2 genes in different animals provide a good example in demonstrating correlation between the genome size and the size of intronic regions (Table 1). With the exception of the mouse and X. genes, decrease in Ace2 gene density is due to the extraordinary large first intron. Ten out of eighteen Ace2 exons show length variations in the six animals, however the exon boundaries in all six genes are identical as indicated by arrows in Supplementary data 2. Although the zebrafish genomic Ace2 sequence has not yet been completed, it is reasonable to predict that the exon boundaries of zebrafish Ace2 are similar to those of the other Ace2s. Disulfide bond formations between cysteine residues are important for the functions of many proteins. We found seven cysteine residues conserved in all ACE2s (Supplementary data 2) suggesting that they may be responsible for maintaining the conformation, stability (Maiti and Surewicz, 2001) or biological activities (Guo et al., 2004). ACE2 residues Arg 273 (critical for substrate binding), His 345 and His 505, involved in catalysis, as proven by site-directed mutagenesis (Guy et al., 2005), are conserved in all 7 genes. Glycosylation of mammalian ACE2 is heterogeneous and differs in terms of extent of glycosylation, sites of glycosylation and structures of attached oligosaccharide units (Ripka et al., 1993). Rabbit testicular ACE (tACE) has 5 potential N-glycosylation sites and the first 2 sites are sufficient and essential for normal synthesis and processing of the active enzyme (Sadhukhan and Sen, 1996), whereas deglycosylated human tACE retains full enzymatic activity (Yu et al., 1997). We scanned the potential N-glycosylation sites from the 7 ACE2 sequences and found 7 sites in human ACE2, 5 in mouse ACE2, 11 in chicken ACE2, 3 in Xenopus ACE2, 11 in Fugu ACE2, 9 in Tetraodon ACE2 and 4 in zebrafish ACE2. There was only one site in human ACE2 (Asn53), which was conserved in other ACE2s except that of zebrafish, thus suggesting an important functional role for the putative N-glycosylation at this site. With the recent SARS pandemic and a possible future bird-flu pandemic looming, the transmission of pathogens from animals to human has received intense scientific attention. A study of the possible transmission of SARS-CoV in domestic poultry suggests that domestic poultry is unlikely to have been the reservoir, or associated with dissemination, of SARS-CoV (Swayne et al., 2004). However, is it possible that birds can become a reservoir? We tried to answer this question by studying the ACE2 receptor binding domain. It has been reported that amino acid residues 31, 41, 353, 355 and 357 of mammalian ACE2 are responsible for the binding with SARS-CoV S protein (Li et al., 2005). Comparing these binding sites from human, palm civet and chicken ACE2s showed that only residue 31 is variable (Lys in human, Thr in palm civet, and Glu in chicken). Interestingly, where palm civet ACE2 efficiently binds the S protein from SARS-CoVs isolated during the 2002–2003 SARS outbreak, the much less severe 2003–2004 outbreak and from the palm civet, human ACE2 could only bind efficiently to the S protein from the earlier human outbreak (2002–2003). This probably indicates that residue 31 plays an important role in the binding between ACE2 and SARS-CoV S protein. Therefore, since chicken ACE2 contains a Glu at position 31 (Supplementary data 2), we hypothesize that this may be an obstacle in the binding between SARS-CoV S protein and chicken ACE2.

The synteny of Ace2–Nhs loci was highly conserved between mammalian and bird, although their evolution diverged from Amniota 300 million years ago. Amniota and Amphibia have diverged from Sarcocryptopygi for 350 million years, and as a result of evolution the locus of Amphibia could either contain many other genes or have been rearranged. This was deduced from the 3 Xenopus scaffolds shown in Fig. 1, which have other genes in the flanking regions. Although there are gaps in the X. tropicalis genome assembly, the syntenies of genes between Ace2 and Nhs appear to be conserved. After 450 million years of evolution the loci are quite different between Actinopterygii and Sarcocryptopygi, i.e. there was no gene in the intergenic region between Ace2 and Nhs in pufferfish (Actinopterygii) whereas at least 8 genes were in the region of Sarcocryptopygi. It is unknown whether genomic rearrangement removed genes from the fish loci or whether they were inserted into Sarcocryptopygi locus. An ongoing genome project (Venkatesh et al., 2005) for a small genome (1200 Mb) cartilaginous fish (elephant fish), the ancestor of Sarcocryptopygi and Actinopterygii, may provide answers to this question.

In order to locate the promoters of human and Xenopus Ace2, we ligated the 1.2 kb (human) and 1 kb (Xenopus) 5′UTRs upstream of the reporter EGFP. Both constructs drove the expression of EGFP in the heart of transgenic Xenopus and in transfected Vero E6 cells (data not shown). The reduced promoter sizes of 0.6 kb (human) and 0.5 kb (Xenopus) still maintains tissue specificity. The 2.8 kb fragment from the intergenic region between Tetraodon Ace2 and Nhs genes is able to drive the EGFP expression too, as in the case of the human and Xenopus promoters. NHS is an X-linked disorder characterized by congenital cataracts, dental anomalies, dysmorphic features, and, in some cases, mental retardation (Burdon et al., 2003), therefore the gene may be expressed during eye, teeth and brain developments. Nhs and Ace2 genes are in a head-to-head orientation in the Tetraodon locus and expression driven by Nhs promoter was expected to be in the eyes and brain regions. However, we were unable to detect the reporter in transgenic Xenopus using the reverse orientation of the 2.8 kb fragment. In mammals, the kidney is another main organ expressing ACE2, but in our transgenic studies, the reporter was clearly absent from the kidney. The expression could only be observed on the heart surface, which may correspond to the expression in endothelial cells of the arteries, arterioles and venules (Oudit et al., 2003). As reported previously, detection methods with higher sensitivity such as quantitative RT-PCR (Hammer et al., 2002) could be used to identify gene expression in other tissues. Even so, the expected site of significant expression levels should still be the heart tissue. We also observed that the expression of all three constructs was shifting from ventricle to the truncus arteriosis or the entire heart but very seldom to atria alone. Transgenic frog lines are being established, and these may be useful for the secondary screening of drugs that target the up-regulation of ACE2 expression in heart blood vessels. Although the transgenic studies have shown tissue specificity and some potential GATA transcription factor binding sites have been found in the putative promoters of Ace2 genes in 7 different vertebrates, the identification of tissue specific elements requires further studies, either using cellular transfection with
suitable heart endothelial cell lines or using transgenic animal models with mutated promoters. Since previous studies have implicated members of the GATA family in endothelial tissue development (GATA-2/3), as well as cardiac function (GATA-4/5/6; Jiang and Evans, 1996), we speculated that the GATA motifs identified in this study could be crucial for the transcription of "Ace2 in cardiac tissues. The conceptual amino acid sequence of exon 18 of "Xenopus ACE2 was aligned with 6 other counterparts using all three reading frames. However, only the ORF encoding for the large cytoplasmic domain observed in "X. tropicalis is typical of amphibian ACE2, the gene from "X. laevis may be useful for additional characterization. In this paper we report five ACE2s from non-mammalian vertebrates. Several lines of evidence, such as identical exon–intron structure, conserved synteny, sequence identity, similar protein primary structure, sequences along the phylogeny and tissue specificity of putative promoters, showed that all the identified "Ace2 genes were the orthologs of mammalian "Ace2.

Acknowledgements

We thank the IMCB DNA Sequencing Facility for its help in sequencing. This project was supported by the Agency for Science and Technology (A*STAR) Singapore.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2006.03.010.

References

Aparicio, S., et al., 2002. Whole-genome shotgun assembly and analysis of the genome of "Fugu rubripes. Science 297, 1301–1310.
Burdon, K.P. et al., 2003. Mutations in a novel gene, NHS, cause the pleiotropic anomalies, and mental retardation. Am. J. Hum. Genet. 73, 1120–1130.
Burrell, L.M., Johnston, C.L., Tikellis, C., Cooper, M.E., 2004. ACE2, a new regulator of the renin-angiotensin system. Trends Endocrinol. Metab. 15, 166–169.
Chou, C.F., Tohari, S., Brenner, S., Venkatesh, B., 2004. Erythropoietin gene from a teleost fish, "Fugu rubripes. Blood 104, 1498–1503.
Chou, C.F., et al., 2005. A novel cell-based binding assay system reconstituting interaction between SARS-CoV S protein and its cellular receptor. J. Virol. Methods 123, 41–48.
Christoffels, A., Koh, E.G., Chia, J.M., Brenner, S., Aparicio, S., Venkatesh, B., 2004. "Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Mol. Biol. Evol. 21, 1146–1151.
Crackower, M.A., et al., 2002. Angiotensin-converting enzyme 2 is an essential regulator of heart function. Nature 417, 822–828.
Crollius, H.R., et al., 2000. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis. Genome Res. 10, 939–949.
Donoghue, M., et al., 2000. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1–9. Circ. Res. 87, E1–E9.
Donoghue, M., et al., 2003. Heart block, ventricular tachycardia, and sudden death in "Ace2 transgenic mice with downregulated connexins. J. Mol. Cell. Cardiol. 35, 1043–1053.

Douglas, G.C., et al., 2004. The novel angiotensin-converting enzyme (ACE) homolog, ACE2, is selectively expressed by adult Leydig cells of the testis. Endocrinology 145, 4703–4711.
Guo, Z.Y., Jia, X.Y., Feng, Y.M., 2004. Replacement of the interchain disulfide bridge-forming amino acids A7 and B7 by glutamate impairs the structure and activity of insulin. Biol. Chem. 385, 1171–1175.
Guy, J.L., Jackson, R.M., Jensen, H.A., Hooper, N.M., Turner, A.J., 2005. Identification of critical active-site residues in angiotensin-converting enzyme-2 (ACE2) by site-directed mutagenesis. FEBS J. 272, 3512–3520.
Hamming, I., Timens, W., Bultuijs, M.L., Lely, A.T., Navis, G.J., van Goor, H., 2004. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. J. Pathol. 203, 631–637.
Harmer, D., Gilbert, M., Borman, R., Clark, K.L., 2002. Quantitative mRNA expression profiling of ACE2, a novel homologue of angiotensin converting enzyme. FEBS Lett. 532, 107–110.
Ikonomi, P., Noguchi, C.T., Miller, W., Kassahun, H., Hardison, R., Schechter, A.N., 2000. Levels of GATA-1/GATA-2 transcription factors modulate expression of embryonic and fetal hemoglobins. Gene 261, 277-287.
Imai, Y., et al., 2005. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature 436, 112–116.
Jaillon, O., et al., 2004. Genome duplication in the teleost fish "Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431, 946–957.
Jiang, Y., Evans, T., 1996. The Xenopus GATA-4/5/6 genes are associated with cardiac specification and can regulate cardiac-specific transcription during embryogenesis. Dev. Biol. 174, 258–270.
Kroll, K.L., Amaya, E., 1996. Transgenic "Xenopus embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. Development 122, 3173–3183.
Kuba, K., et al., 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. Nat. Med. 11, 875–879.
Li, W., et al., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426, 450–454.
Li, W., et al., 2005. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J. 24, 1634–1643.
Lohr, J.L., Yost, H.J., 2000. Vertebrate model systems in the study of early heart development: "Xenopus and zebrafish. Am. J. Med. Genet. 97, 248–257.
Maiti, N.R., Surewicz, W.K., 2001. The role of disulfide bridge in the folding and stability of the recombinant human prion protein. J. Biol. Chem. 276, 2427–2431.
Menard, J., Patchett, A.A., 2001. Angiotensin-converting enzyme inhibitors. Adv. Protein Chem. 56, 13–75.
Oudit, G.Y., Crackower, M.A., Backx, P.H., Penninger, J.M., 2003. The role of ACE2 in cardiovascular physiology. Trends Cardiovasc. Med. 13, 93–101.
Ripka, J.E., Ryan, J.W., Valido, F.A., Chung, A.Y., Peterson, C.M., Urry, R.L., 1993. N-glycosylation of forms of angiotensin converting enzyme from four mammalian species. Biochem. Biophys. Res. Commun. 196, 503–508.
Sadhukhan, R., Sen, I., 1996. Different glycosylation requirements for the synthesis of enzymatically active angiotensin-converting enzyme in mammalian cells and yeast. J. Biol. Chem. 271, 6429–6434.
Skegg Jr., L.T., Kahn, J.R., Shumway, N.P., 1956. The preparation and function of the hypertensin-converting enzyme. J. Exp. Med. 103, 295–307.