Two novel DNA motifs are essential for BACE1 gene transcription

Yan Xiang1*, Shasha Meng1*, Jinfeng Wang1, Songyang Li1, Jingru Liu1, Hongmei Li1, Tingyu Li1, Weihong Song1,2 & Weihui Zhou1

1Ministry of Education Key Laboratory of Child Development and Disorders; Chongqing Key Laboratory of Translational Medical Research in Cognitive Development and Learning and Memory Disorders, Children’s Hospital of Chongqing Medical University, 136 ZhongshanEr Lu, Yuzhong District, Chongqing 400014, China, 2Townsend Family Laboratories, Department of Psychiatry, Brain Research Center, University of British Columbia, Vancouver, British Columbia, V6T1Z3, Canada.

BACE1 gene encodes for β-Site amyloid β precursor protein (APP)-cleaving enzyme1, which is required for generating amyloid β protein (Aβ). Deposition of Aβ in brain plays an essential role in Alzheimer’s Disease (AD) pathogenesis. BACE1 gene has a tissue-specific expression pattern and its expression is tightly regulated at transcriptional level. Core promoter is a minimal DNA sequence to direct transcription initiation and serves as a converging platform for the vast network of regulatory events. Here we identified the core promoter of human BACE1 gene, which is a 71 nucleotides region absent of typical known core promoter elements and is sufficient to initiate a basal transcription. Two novel DNA motifs, designated TCE1 and TCE2, were found to be involved in activating the transcription of human BACE1 gene in a synergistic way. Two single nucleotide mutations in these motifs completely abolished the promoter activity. In conclusion, our studies have demonstrated that novel DNA motif TCE1 and TCE2 in human BACE1 gene promoter are two essential cis-acting elements for BACE1 gene transcription. Studies on how these two motifs being regulated by different stimuli could provide insights into the molecular mechanisms underlying AD pathogenesis and pharmaceutical potentials of targeting these motifs for AD treatment.

Amyloid β protein (Aβ) is the major component of amyloid neuritic plaques in the brain of AD and its deposition in the brain plays an essential role in AD pathogenesis1. Aβ is generated from APP by β- and γ-secretase cleavages2. BACE1 gene encodes for β-Site APP cleaving enzyme1 (BACE1, also known as Asp2 or memapsin2), a type 1 membrane-associated aspartyl protease of 501 amino acids3–6. BACE1 is a β-secretase in vitro7–11. BACE1 cleaves APP at two β-secretase sites, at Asp1 site to produce CTFβ C99 in amyloidogenic pathway, or at Glu11 site to produce CTFβ C89 in non-amyloidogenic pathway12. CTFβ C99 is the substrate of presenilin-dependent γ-secretase complex to produce Aβ, thus cleavage of APP by BACE1 at Asp1 site is essential in the amyloidogenic pathway. The β-cleavage is considered a rate limiting step for Aβ production. Degradation of BACE1 is mediated by the ubiquitin proteasome pathway13–15. A slight increase of BACE1 expression is sufficient to promote C99 and Aβ production under pathological conditions16. Reduction of Aβ by inhibition of BACE1 activity is considered to have major pharmaceutical potential for AD therapy17,18.

The BACE1 gene has a tissue and stimuli-specific expression pattern19,20. BACE1 mRNA has been found mainly in neuronal cells and in the pancreas21,22. BACE1 expression is tightly regulated at the level of transcription19. The upstream sequence of the BACE1 gene contains many putative transcription factor binding sites23–25. Several transcription factors (TFs) have been experimentally confirmed to bind to the regulatory region of BACE1 gene and to positively regulate its expression. Transcription factor specificity protein 1 (Sp1) was the first one discovered to play an important role in regulating human BACE1 expression19. Hypoxia inducible factor 1 (HIF-1) was found to bind to a hypoxia responsive element (HRE) on BACE1 promoter and mediated the increases of BACE1 transcription induced by hypoxia22–25. Signal transducer and activator of transcription 1 (STAT1) is responsible for BACE1 constitutively expression in neurons and inducible expression by interferon γ (IFNγ) in astrocytes26,27. Transcriptional factor Yin Yang 1 (YY1) and nuclear factor of activated T-cells 1 (NFAT1) were also reported to up-regulating BACE1 transcription28–30. There are also many TFs and their binding sites down-regulating BACE1 gene transcription. Two GATA-1 sites located at -106 to -97 and -209 to -200 from authentic ATG codon of rat BACE1 were found to act as suppressors of BACE1 transcription31. The peroxisome proliferator-activated receptor-γ (PPARγ) can bind to PPARγ responsive element (PPRE) of rat BACE1 promoter and suppress BACE1 transcription, and decreasing PPARγ level by proinflammatory cytokines is responsible for activation of BACE1 promoter32,33. Functional binding sites of Nuclear factor κB (NFκB) were identified in...
BACE1 promoter region\textsuperscript{11,33,34}. NFkB was reported to differentially regulate BACE1 gene transcription, which depends on cell-type or cell-condition\textsuperscript{23,33,34}. Under physical condition or in neuronal and non-activated astrocytic cells activated NFkB inhibits BACE1 gene expression, while it activates BACE1 in higher AB concentrations or in activated astrocytes. The majority of cis-acting elements of BACE1 gene promoter are located far from transcriptional start site (TSS). There is little research on the function and structure of the proximal region of BACE1 promoter or core promoter near TSS.

Core promoter is a minimal DNA sequence to direct transcription initiation and serves as a converging platform for the vast network of regulatory events by interaction of transcription factors and proximal or distal regulatory DNA motifs in the promoter\textsuperscript{25,26}. Based on the distribution of TSS, core promoter was divided into two classes as focused (sharp) and dispersed (broad) promoter\textsuperscript{35,36}. In general, a focused promoter is associated with genes containing typical core promoter elements (TATA, Inr, DPE or MTE), lacks positioned nucleosomes in the promoter region and has a labile transcriptional state and a low GC content; whereas a dispersed promoter correlates with housekeeping genes with positioned nucleosomes in the promoter region, and it does not have known core promoter elements, and has a stable transcriptional state and a high GC content\textsuperscript{37,38}. Analysis of focused core promoter in RNA polymerase II (RNAPII)-transcribed genes suggests that a core promoter typically spans from around nucleotide ~40 to +40 relative to transcription start site and consists of functional sub-regions termed core promoter elements such as Inr, TATA box, BRE, DPE, MTE, XCPE1. Recent evidence indicates that different core promoter type can interact with different transcription components and such diversity plays an important role in driving cell-specific transcription and regulatory events\textsuperscript{39}. BACE1 gene promoter lacks the typical CAAT and TATA box but within 2 kb upstream region of the TSS\textsuperscript{40}. In-depth analysis of structure and function of BACE1 core promoter will greatly advance our understanding the transcriptional mechanisms of TATA-less core promoter and may benefit for developing new strategies in inhibiting BACE1 expression in prevention and treatment of Alzheimer’s disease.

In this study, we defined the boundary of the proximal region of human BACE1 promoter, which is a 104 base pairs in length and sufficient for high levels of expression in cells. We identified the core promoter region of 71 nucleotides from −550 to −480 bp relative to ATG of the BACE1 gene and this core promoter region is able to initiate a basal transcription of downstream gene. BACE1 gene core promoter lacks known core promoter elements. Two novel DNA motifs, designated as TCE1 and TCE2, were found to be involved in activating the transcription of BACE1 gene in a synergistic way. Reducing the spacing between TCE1 and TCE2 decrease the promoter activity and two single nucleotide mutations in these motifs are able to completely abolish promoter activity of BACE1 gene. Collectively, our studies have demonstrated that novel DNA motif TCE1 and TCE2 identified in human BACE1 gene promoter are essential for controlling transcription of BACE1 gene.

**Results**

**Determination of the boundary of BACE1 minimal and core promoter.** Our previous study demonstrated that the region from −1311 to −400 bp upstream of the translation start codon (as +1) of the human BACE1 gene has promoter activity\textsuperscript{41}. In order to identify core promoter, first we constructed luciferase reporter plasmid p3BU-1273/-400 containing sequence −1273 to −400 bp region upstream of translating start site of BACE1 gene inserted in pGL3 Basic vector. DNA sequencing confirmed that the sequence of inserted BACE1−1273 to −400 in p3BU-1273/-400 is identical to 117187784-117186911 of chromosome 11 (GRCh37/hg19), except for one SNP rs536817 (chr11:117187441), which is G in our sequence. The pGL3-Basic vector lacks eukaryotic promoter and enhancer sequences upstream of a reporter luciferase gene. Luciferase activity in cells transfected with reporter plasmid depends on the expression of luciferase gene, and the expression level depends on insertion and proper orientation of a functional promoter upstream from the reporter luciferase gene.

To identify the boundary of minimal 5’ flanking region required for BACE1 promoter activity, several deletion plasmids were generated and transfected into HEK293 cell. Luciferase activity measured from cells transfected with the original pGL3-Basic vector served as control. The results of luciferase assay indicated that p3BU-1273/-400, p3BU-1149/-400, p3BU-919/-400, p3BU-800/-400, p3BU-700/-400, p3BU-583/-400 have similar promoter activity and are significantly higher than pGL3-Basic control (Fig. 1a, b). Further deletion of 83 bp (p3BU-500/-400) almost abolished luciferase activity. This indicates that the 184 bp fragment (−583/−400) contains the functional promoter region and the 83 bp long region between −583 to -501 contains important cis-acting elements required for human BACE1 gene transcription. To narrow down this important region, more deletion plasmids were constructed and luciferase assay (Fig. 1c, d) showed that further deleting 10 bp from 5’end of -583/−400 drastically reduced luciferase activity, but is still higher than negative control. Additional deletion of 14 bp (p3BU-560/-400) displayed similar luciferase activity as p3BU-574/-400. However, continuous deletion of 20 bp (p3BU-540/-400) caused significant decrease of luciferase activity, which is close to the levels of pGL3-Basic control. The results indicate that region -583 to -574 and -560 to -540 have important and functional elements required for BACE1 promoter activity.

Since the region -583 to -400 has significant promoter activity and further deletion from 5’ end causes drastically decrease of activity, p3BU-583/-400 was chosen as a template to investigate the effect of 3’ end on promoter activity. A series of 3’ end deletion plasmids were constructed and transfected into HEK293 cell. Luciferase assay showed that deleting two 20 bp from 3’end of -583/-400 (resulting in p3BU-583/-420 and p3BU-583/-440, respectively) had no significant effects on luciferase activity (Fig. 1e). Further deletion of 20 bp (p3BU-583/-460) caused significant decrease of luciferase activity compared to p3BU-583/-440. However, deleting additional 20 bp from 3’end of -583/-460 (p3BU-583/-480) surprisingly recovered luciferase activity to almost same as p3BU-583/-400. Further deleting 20 bp from 3’end of -583/-480 significantly reduced luciferase activity and continuous deletion of another 10 bp (p3BU-583/-510) caused further decrease of luciferase activity and is closed to the levels of pGL3-Basic control (Fig. 1e). It suggests that −480 is the 3’ boundary of region required for BACE1 minimal promoter activity and the region of -510 to -480 is essential for BACE1 promoter activity.

To validate whether the position of nucleotide -480 is the 3’ boundary of region required for BACE1 minimal promoter activity in the different context of 5’ end, more 5’ end deletion plasmids were generated and transfected into HEK293 cell. Luciferase activity showed that a series of plasmids ended with -480 as 3’ end have similar pattern of luciferase activity as a series of plasmids 3’ end with -400 (Fig. 1f). Deleting 14 bp from 5’ end of BACE1 fragment -583 to -480 drastically reduced luciferase activity, further deleting two 10 bp region (p3BU-560/-480 and p3BU-550/-480) showed similar activity as p3BU-570/-480 and continuing to delete 10 bp from 5’end of -550 to -480 caused further significant decrease of luciferase activity in p3BU-540/-480. It suggested that 71 bp region of -550 to -480 still has significant promoter activity. Inserting the 104 bp fragment of -583 to -840 bp in reverse direction in plasmid p3BU-583/-480R abolished luciferase activity (Fig. 1f). Taken together, our data demonstrated that this 104 nucleotides region from -583 to -480 bp upstream of the translational start site is sufficient for high level expression of BACE1 gene and is the BACE1 minimal promoter, and furthermore, the 71 nucleotides region of -550/-480 bp functions as BACE1 core
promoter, which is responsible for determining the basal transcription of BACE1 gene.

Identification of the cis-acting regions. Our study suggests that there are a few sub-regions in BACE1 minimal promoter sequence important for BACE1 gene transcription, especially the region -583 to -574 and -510 to -480. To further investigate whether these regions are functional, reporter plasmids with random single nucleotide mutation were constructed with plasmid p3BU-583/-400 as template and the mutant plasmids were transfected into HEK293 cell. Luciferase assay showed that the single nucleotide mutation from C to T at -580 (-580C/T) had no negative effects, but mutation -578G/A, -510A/C or -482T/A negatively affected luciferase activity of plasmid p3BU-583/-400 (Fig. 2a and b).

Since position -580 and -578 are very close to the vector sequence in plasmid p3BU-583/-400M580T and p3BU-583/-400M578A, in order to exclude the possible effects of flanking vector sequence on luciferase activity of inserted fragment and verify the effects of those mutation, plasmid p3BU-1149/-400 carrying the mutations with extended 5' end were generated (Fig. 2c). Luciferase assay showed that mutations 578G/A, 510A/C still significantly reduced the luciferase activity of wildtype plasmid p3BU-1149/-400 (Fig. 2d). Next, the effects of surrounding nucleotides of nucleotide -578G and -510A were investigated to determine if there are cis-acting regions in the vicinity of these two nucleotides. Several nucleotide cluster mutation plasmids were constructed as shown in Fig. 2e. Luciferase assay demonstrated that five cluster mutations near nucleotide -578G and -510A significantly reduced the luciferase activity of wildtype plasmid p3BU-1149/-400 (Fig. 2f) and suggested that there are functional cis-acting regions near -578G or -510A, respectively.

The cis-acting regions have a synergistic effect on activating BACE1 transcription. To investigate whether the afore-mentioned different regions have a synergistic effect on BACE1 promoter activity, reporter plasmids containing single nucleotide mutations in both regions were constructed with plasmid p3BU-583/-400 as template and transfected into HEK293 cell. Compared with single nucleotide mutation of 578G/A (p3BU-583/-400M578A) or 580C/T...
double nucleotides mutation of 578G/A and 580C/T (p3BU-583/-400M580T578A) showed similar luciferase activity as single nucleotide mutation of 578G/A, suggesting that mutation of 580C/T has no additional effects on luciferase activity of plasmid p3BU-583/-400M578A (Fig. 3a,b).

However, luciferase assay showed that double mutations 578G/A and 482T/A (p3BU-583/-400M578A482A) showed similar luciferase activity as single nucleotide mutation of 578G/A, suggesting that mutation of 482T/A had no effects on promoter activity. The luciferase assay of the double mutant 510A/C and 482T/A (p3BU-1149/-400M510C482A) also suggest that the mutation 482T/A had no significant effects on \( \text{BACE1} \) gene transcription.

Identification of the nucleotide composition for the two cis-acting elements. In order to identify possible proteins binding to surrounding sequences of nucleotide -578G or -510A, a 21-nucleotide sequence corresponding to 10 nucleotides flanking each side of -578G or -510A was used for transcriptional factor analysis using MatInspector\(^4\). Only one protein Myc-interacting Zn finger protein 1 (MIZ1) was found to match the sequence in \( \text{BACE1} \)-588/-568 (CACCTCGGCAGAGGGCATCCC). The MIZ1-binding protein was shown to bind to this sequence in a ChIP assay. As expected, plasmid p3BU-1149/-400M578A510C showed much weaker luciferase activity than p3BU-1149/-400EGFP (Fig. 3e). These results confirmed that combining two single nucleotide mutations of 578A and 510C had a synergistic effect on \( \text{BACE1} \) promoter activity. The data suggest that the two cis-acting regions near nucleotides -578G and -510A play a pivotal role and function cooperatively in determining \( \text{BACE1} \) gene transcription.

Next, we examined the possible effects of flanking vector sequence on luciferase activity. Compared with single nucleotide mutation of 578G/A (p3BU-1149/-400M578A) or 510A/C (p3BU-1149/-400M510C), double mutant 578G/A and 510A/C (p3BU-1149/-400M578A510C) drastically reduced luciferase activity of wildtype plasmid p3BU-1149/-400 and had much lower luciferase activity than each single mutant 578G/A or 510A/C (Fig. 3c, d). To further verify the significant effects of the double mutation of 578G/A and 510A/C on \( \text{BACE1} \) gene transcription, four plasmids with green fluorescent protein (EGFP) as a reporter gene were constructed by replacing the luciferase reporter gene with EGFP gene in plasmids pGL3-Basic, pGL3-Promoter, p3BU-1149/-400 and p3BU-1149/-400M578A510C. Empty control plasmid pGL3basic-EGFP didn’t show florescence 24 hours after transfection into HEK293 cell, positive plasmid pGL3promoter-EGFP, containing a strong SV-40 promoter, showed very strong florescence. Plasmid p3BU-1149/-400EGFP displayed significant brighter florescence than pGL3basicEGFP. As expected, plasmid p3BU-1149/-400M578A510C EGFP displayed much weaker florescence than p3BU-1149/-400EGFP (Fig. 3e). These results confirmed that combining two single nucleotide mutations of 578A and 510C had a synergistic effect on \( \text{BACE1} \) promoter activity. The data suggest that the two cis-acting regions near nucleotides -578G and -510A play a pivotal role and function cooperatively in determining \( \text{BACE1} \) gene transcription.
sequence identified by MatInspector is gatgcCCTCtg (capital letter represents core sequence), which is located in antisense strand of \textit{BACE1} -580 to -570 (CAGAGGGCATC). Since the candidate binding site of MIZ1 is located in antisense strand of \textit{BACE1} gene, a reporter plasmid (p3BU-1149/-400MIZ) with candidate binding site of MIZ1 at sense strand was constructed and transfected into HEK cells. The luciferase activity of p3BU-1149/-400MIZ was significantly lower than wildtype p3BU-1149/-400 (Fig. 4a, b), suggesting that candidate MIZ1 binding site identified by MatInspector may not be functional. Searching transcriptional factors binding to \textit{BACE1} gene promoter sequence of -520/-500 (TACACTTCCCAGCGATCCCAG) with program MatInspector resulted in no hit.

Since there is no known transcriptional factor bound to these two important regions, next we defined the boundaries of both regions and investigated the effects of nucleotide substitution on promoter activity. We first generated p3BU-1149/-400 mutant plasmids carrying mutations surrounding -578G or -510A, and each nucleotide was replaced with three other nucleotides, respectively. The luciferase assay results showed that the majority of nucleotide replacements in region between -583 and -573 had negative effects on promoter activity (Fig. 4c), indicating that the region between -583 and -573 contains a positive-regulatory \textit{cis}-acting element and the sequence of CGGCAGAGGGC was named as \textit{BACE1} Transcription Critical Element 1 (TCE1), which may bind to unknown transcriptional factor. The nucleotide replacement at -584T with C and at -572A with G even increased the promoter activity. The nucleotide replacement at -580C with T also significantly increase the promoter activity, suggesting that the wildtype sequence of TCE1 may not be the ideal sequence for binding with a transcriptional factor.

For the functional region surrounding -510A, the boundary seems to be -515 and -509, since almost every replacement between -515 and -509 significantly reduce the promoter activity (Fig. 4d). However, the replacement at -516C with three other nucleotides and the majority of nucleotide replacement between -508 and -505 had no negative effects. The region between -515 and -509 contains a positive-regulatory \textit{cis}-acting element and the sequence of TTCCCAG was named as \textit{BACE1} Transcription Critical Element 2 (TCE2). Comparing the effects between TCE1 and TCE2 on \textit{BACE1} promoter activity, many nucleotide replacements in TCE2 caused 70% reduction and nucleotide replacements in TCE1 only caused 40-50% reduction, suggesting that TCE2 has stronger effects than TCE1 on \textit{BACE1} transcription.

\textbf{TCE1 is highly conserved and serves as a promoter proximal element.} Next we analysed evolutionary conservation of 100 vertebrate species within the region of -583/-480. Three highly conserved regions were identified and they are -583/-570 (chr11:117187094-117187081, CGGCAGAGGGCATC), -551/-544 (chr11:117187062-117187055, CCGGAAGC) and -528/-509
The box LUC with large arrowhead represents the coding sequence of the luciferase reporter gene and transcriptional direction. The numbers indicate the endpoints of each construct, with +1 as the adenine of the physiological translational start codon of the BACE1 gene.

Luciferase activity was measured 24 h after transfection, and Renilla luciferase activity was used to normalize for transfection efficiency. Data are shown as percentage of control samples transfected with wild-type plasmid. The values represent means ± standard error of the mean (n = 3–6). # or * P < 0.05 by analysis of variance (ANOVA) with the post-hoc Newman-Keuls test when comparing with control. (d) (f) Schematic diagram of plasmid constructs containing each nucleotide substitution in plasmid p3BU-1149/-400. Top row indicates wild-type nucleotide and bottom row indicates the nucleotide position.

Figure 4 | Identification of two transcription critical elements for BACE1 gene. (a) Schematic diagram of plasmid constructs. The two horizontal lines indicate two DNA strands of the BACE1 promoter region; inverted letter of MIZ1- indicate its original orientation at minus strand; small arrow indicates MIZ1 motif. (chr11:117187039-117187020, GGGAGACTACACTTCCCAG) (Fig. 5a). Nucleotide base level alignment analysis of five species, human and baboon from primate subset, mouse and rat from Euarchontoglires subset and elephant from Afrotheria subset, revealed that sequence motifs of regions -583/-569, -551/-544 and -528/-509 are exactly same between five species, although the single nucleotide of -582 (chr11:117187093) and -577 (chr11:117187088) in region -583/-569 is not conserved from 100 vertebrate species (Fig. 5a). The region of -570/-550 is not highly conserved between five species, which suggests that the sequence of -570/-550 region may have little effect on BACE1 transcription.

TCE1 is located immediate upstream of the core promoter. To investigate whether it is orientation-dependent and activates other gene promoter, we constructed reporter plasmids p3BU-1149/-400TCE1R and pGL3p-TCE1. The TCE1 sequence in p3BU-1149/-400 was replaced with the sequence of reverse direction from CGGCAGAGGGC to GCCCTCTGCCG to generate p3BU-1149/-400TCE1R, and the TCE1 sequence was inserted in front of SV40 promoter in plasmid pGL3-Promoter to generate pGL3p-TCE1 (Fig. 5b,d). The luciferase activity of p3BU-1149/-400TCE1R was less than half of the control’s activity, suggesting that TCE1 is orientation-dependent (Fig. 5c). However, placing the TCE1 sequence in front of SV40 promoter had no effect on luciferase activity of enhancer-less reporter plasmid pGL3-Promoter and it suggests that the function of TCE1 is specific to BACE1 core promoter (Fig. 5e).

The distance between two cis-acting elements is important for gene transcription in some cases. Since deletion of the region -570 to -550 had no effects on BACE1 basal transcription, the effect of the space between TCE1 and core promoter on BACE1 transcription was investigated to evaluate whether deleting region -570 to -550 had an effects on synergistic action of TCE1 and core promoter. Region -570 to -550 reverse mutation plasmid p3BU-1149/-400M570-551R and deletion mutation plasmid p3BU-1149/-400M570-551D were constructed with p3BU-1149/-400 as template (Fig. 5f). The sequence CCCAGACCCCTCTCCAGCCC of -570/-551 was changed to CCCCGACCTCTCCCCAGACCC in p3BU-1149/-400M570-551R or deleted in p3BU-1149/-400M570-551D. Luciferase assay revealed that reverse mutation of region -570/-550 in p3BU-1149/-400M570-551R had no effects on luciferase activity and deletion mutation of region -570/-550 in p3BU-1149/-400M570-551D resulted in significant decrease in promoter activity (Fig. 5g). Taken together, TCE1 is an orientation-and position-dependent cis-acting promoter proximal element and a specific integral part of the BACE1 core promoter.

BACE1 core promoter does not contain typical core promoter elements. The common core promoter elements include the Inr (initiator, yyanwyy), the TATA box (tatawaar), BRE (TFIIB recognition element, ssrccgcc), DPE (downstream core promoter element,rgwyvt), MTE (motif ten element,csarcssaacgs), and XCPE1(X core promoter element 1, dsgygrrasm)42. Sequence analysis reveals that BACE1 core gene core promoter region of -583/-400 has no TATA box, BRE, DPE,MTE, or XCPE1. However, there are two putative Inr elements. Inr (initiator) is a recognition site for the binding of transcriptional factor TFIIID and has a consensus sequence of YYANWYY in human. Two putative Inr elements are located at -569/-563(CCGAGACC) and -521/
-515(CTACACT) respectively. Because upstream putative Inr site -569/-563 is located between -570 and -550 in which the sequence are not functional as shown before, and deleting region -570/-550 has no effects on BACE1 basal transcription, upstream putative Inr is most likely not functional. Downstream putative Inr site is located at -521/-515 and is not consistent with published and reference gene TSS at GenBank, which are located between nucleotide 453 to 461 upstream of translational start site ATG. To verify whether it is functional Inr site, first we constructed plasmid p3BU-1149/-400InrM by replacing Inr site with an reverse Inr sequence (-521/-515 CTACACT was changed to tcACAtc) (Fig. 6a). The luciferase activity of reporter plasmid p3BU-1149/-400InrM was about 30% of the wildtype p3BU-1149/-400 (Fig. 6b), even though mutated sequence tcACAtc was still corresponding to Inr consensus sequence yyanwyy. Secondly we performed 5'-RACE experiment to define the actual TSS in reporter plasmid p3BU-1149/-400InrM and it is very close to known TSS of -461 in reference gene (NM_012104.4) (Fig. 6c). Because of technical limitation, 5'-RACE experiment can only define a few TSS each time. Next we applied dbTSS database to investigate the BACE1 TSS, and the result revealed that BACE1 TSS is not restricted into one nucleotide, but a region around -461. However, there is no tag count for the region -521/-515 (Fig. 6d). Taken together, downstream putative Inr site is not functional as an Inr. It suggests BACE1 core promoter does not contain any known typical core promoter elements.

Discussion

Phylogenetic analysis of DNA sequences is a useful strategy to identify DNA sequence motifs that have been strongly conserved through evolution and well-conserved sequence can provide much stronger evidence of functional relevance. A conservation analysis revealed 92% similarity between rat and mouse and 81% similarity between rat and human within the -600 bp upstream of the authentic ATG codon of BACE1 gene, but similarity between rat and human in the region -1000 bp to -600 bp from ATG drops down to 39% and the similarity is lost beyond -1000 bp21. This observation indicates that the 600 bp region immediate upstream of ATG plays an important role in regulating BACE1 expression, consistent with our results that deleting the upstream region of -583 has no effects on promoter activity. These results also suggest that human and mouse may have
similar regulatory mechanism with 600 bp upstream of the ATG codon and have different regulatory mechanism beyond 600 bp region. Studies on promoter structure and transcriptional regulation of \( \text{BACE1} \) have been mainly focused on the rat and human genes. The cis-acting elements identified and verified experimentally in the rat and human \( \text{BACE1} \) gene need to be carefully considered in term of the location. The ones located over-600 bp from ATG, especially over-1000 bp region may play different roles in the transcriptional regulation of \( \text{BACE1} \) gene between rat and human. The cis-acting elements identified experimentally in human \( \text{BACE1} \) include NF-\( \kappa \)B binding site\(^3\), STAT1 binding site\(^2\), hypoxia responsive element\(^2\), Sp1 binding site\(^1\), NFAT1 binding site\(^3\), and they are located at -3020/-2991, -2311/-2302, -1607/-1602, -1602/-1597, -954/-946 upstream of ATG, respectively. Those confirmed in rat \( \text{BACE1} \) include NF-\( \kappa \)B binding site\(^3\), PPAR\( \gamma \) responsive element\(^1\), YY1 binding site\(^2\), two GATA-1 binding site\(^2\), and they are located at -1521/-1511, -1356/-1338, -575/-560, 209/-200 and -106/-97 from ATG, respectively. Among these elements, only YY1 binding site and GATA-1 binding site are located 600 bp region immediate upstream of authentic ATG. Ge et al. reported a 91 nucleotides region as a minimal promoter at -223/-133 from ATG of human \( \text{BACE1} \) gene\(^4\). Two GATA-1 binding sites and the minimal promoter reported by Ge et al. are actually located at 5’ untranslated region (5’UTR) of mRNA, which have been proved to affect protein translation by a couple of groups\(^4\) and it may affect luciferase translation and activity if 5’UTR include in luciferase reporter plasmid. To focus on examining transcription regulation and to eliminate the interference from post-transcriptional regulation, when constructing reporter plasmids, we intentionally excluded the region of -400 bp to -1. The binding sequence of YY1 in rat \( \text{BACE1} \) promoter is located at -575/-560(ACTCCATCTCGGCAG) upstream of rat authentic ATG, same as the human sequence at -592/-578 (ACTCCACCTCGGCAG) except for one nucleotide difference at -586. Although putative YY1 in human gene overlap partially with TCE1, TCE1 is not like to be YY1 binding site based on following three reasons. First, only six nucleotides out of 15 in the 3’ end of human putative YY1 binding site overlap with 5’ end of -583 bp/578 bp in TCE1. Secondly, deleting 9 nucleotides out of 15 at 5’ end in p3BU-583/-400 has similar luciferase activity as p3BU-700/-400, which include full YY1 binding sequence. It suggests that 5’ end 9 nucleotides of human putative YY1 binding site are not functional. Thirdly, according to reported YYI consensus sequence CGCC-ATnTT\(^4\), core motif of YY1 in rat corresponds to -589/-584 in human and this region does not overlap with TCE1. Therefore, TCE1 is a novel DNA motif to activate transcription of human \( \text{BACE1} \) gene.

RNAPII-transcribed genes typically contain two distinct families of cis-acting transcriptional control sequences, promoters and distal regulatory elements\(^1\). The promoter is the DNA sequence responsible for transcription initiation and can be divided into two sub-

---

**Figure 6 | TCE2 is not an initiator of \( \text{BACE1} \) gene.** (a) Schematic diagram of plasmid constructs. The two horizontal lines indicate two DNA strands of the \( \text{BACE1} \) promoter region; the box LUC with large arrowhead represents the coding sequence of the luciferase reporter gene and transcriptional direction; small arrow indicates tested DNA motif. The numbers indicate the endpoints of each construct, with +1 as the adenine of the physiological translational start codon of the \( \text{BACE1} \) gene. (b) HEK293 cells were transiently transfected reporter plasmids. Firefly luciferase activity was measured 24 h after transfection, and Renilla luciferase activity was used to normalize for transfection efficiency. Data are shown as percentage of control samples transfected with wild-type plasmid. The numbers represent means standard error of the mean (n=6). * P< 0.05 by analysis of t-test when comparing with control. (c) Sequencing result of the 5' RACE product which was cloned into the pGL3-Basic vector. Sequence of the vector, adaptor and \( \text{BACE1} \) 5'UTR are shown underlined. The first nucleotide cytosine next to the 3' of the adaptor is the TSS of human \( \text{BACE1} \) gene. (d) TSS distribution of \( \text{BACE1} \) gene in HEK293 cell from dbTSS(http://dbtss.hgc.jp/). The number under vertical blue bar represent TSS frequency identified from massively parallel sequencing of TSS.
regions of a core promoter and a proximal promoter located immediate upstream of core promoter. Distal regulatory elements include enhancers, silencers, insulators, or locus control regions (LCR). The core promoter is generally defined as the minimal piece of DNA sequence sufficient to guide the assembly of transcription initiation complex at DNA chain and initiate a basal transcription. Typical core promoter spans from around nucleotide −40 to +40 relative to transcription start site and consists of functional DNA motifs termed as core promoter elements. The core promoter elements are short DNA sequence in core promoter and confer distinct properties to the promoter. Our current knowledge of core promoter elements are primarily come from the analysis of focused core promoter in RNAPII-transcribed genes. Although parallel sequencing have been successfully used to study and identified TSS, enhancer and promoter in genome wide,5,12, it is still challenging to define core promoter elements because of its considerable diversity and relatively few study on them, especially in those TATA-less promoters. By in-depth dissection of human BACE1 promoter, our results demonstrated that BACE1 gene core promoter lacks typical known core promoter elements including Inr, TATA box, BRE, DPE, MTE, and XCE1. The newly identified TCE2 from our study is most likely to be a novel unknown core promoter element, which plays an important role in initiating a basal transcription and may be involved in different transcriptional mechanism.

In this study, we defined the boundary of human BACE1 gene minimal and core promoter and found two important DNA motifs, TCE1 and TCE2. TCE1 is likely to be a promoter proximal element and TCE2 is a core promoter element. TCE1 and TCE2 function in a synergistic way and these two novel DNA motifs are essential for transcription of human BACE1 gene. Future study on how these two motifs being affected by different stimuli to regulate BACE1 promoter activity of APP could provide insights into the molecular mechanisms underlying Alzheimer pathogenesis and pharmaceutical potentials of targeting these motifs for AD treatment.

Methods

Cell culture and transfection. Human embryonic kidney 293 (HEK293) cells were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum, 1 mMol/L of sodium pyruvate, 2 mMol/L of L-glutamine and 1% penicillin-streptomycin at 37°C in a humidified incubator containing 5% CO2. Cells were grown to approximately 70% to 80% confluence and plasmid DNA was transfected using Lipofectamine 2000 reagent (Invitrogen) according to the supplier’s protocol, 3.2 μg of total plasmid DNA were transfected into the cells cultured on a 35-mm-diameter dish for 5’-RACE and RNA extraction and 200 ng of total plasmid DNA per well were transfected into the cells cultured on a 48-well plate for luciferase reporter assay.

Plasmid construction. BACE1 promoter luciferase reporter plasmid p3BU-1273/-400 contains human genomic sequence from −1273 to −400 relative to BACE1 gene translating start site as + 1. BACE1 genomic fragment B1U-1273/-400 was amplified with primer B1U1273F (CCGCTCGAGAACCATACCGGC- TTTTGCCTC) /B1U400R (CAACAAGCTTCCAACTTCCAGGCT) and BAC DNA (RP11-970G21) as template. PCR product of B1U-1273/-400 was inserted in front of the firefly luciferase reporter gene in the promoterless reporter plasmid pGL3-Basic (Promega) at Xhol/HindIII sites and inserted sequence was confirmed by DNA sequencing. The methods for constructing other plasmids are described in Supplementary Methods.

Rapid amplification of S’ complementary DNA ends (5’-RACE). The plasmid p3BU-1149/-400 was transfected into the cells and total RNA was extracted from cells with TRI reagent following the manufacturer’s protocol (Sigma). The 5’-end of total plasmid DNA was determined using the dual-luciferase reporter assay system according to the supplier’s protocol (Promega). The firefly luciferase activity was normalized to the Renilla luciferase activity and expressed as relative luciferase units (RLUs) to reflect the promoter activity.

![Scientific Reports](https://www.nature.com/scientificreports/)
25. Li, Q. Y. et al. Salidroside attenuates hypoxia-induced abnormal processing of amyloid precursor protein by decreasing BACE1 expression in SH-SY5Y cells. *Neurosci Lett* 481, 154–158, doi:10.1016/j.neulet.2010.06.076 (2010).

26. Cho, H. J. et al. IFN-gamma-induced BACE1 expression is mediated by activation of JAK2 and ERK1/2 signaling pathways and direct binding of STAT1 to BACE1 promoter in astrocytes. *Glia* 55, 253–262 (2007).

27. Cho, H. J. et al. Constitutive JAK2/STAT1 activation regulates endogenous BACE1 expression in neurons. *Biochim Biophys Acta* 1836, 175–180 (2009).

28. Nowak, K. et al. The transcription factor Yin Yang 1 is an activator of BACE1 expression. *J Neurochem* 96, 1696–1707, doi:10.1111/j.1471-4159.2006.03692.x (2006).

29. Cho, H. J., Jin, S. M., Youn, H. D., Huh, K. & Mook-Jung, I. Disrupted intracellular calcium regulates BACE1 gene expression via nuclear factor of activated T cells 1 (NFAT1) signaling. *Aging cell* 7, 137–147, doi:10.1111/j.1471-4948.2007.00360.x (2008).

30. Cho, H. J. et al. RAGE regulates BACE1 and Abeta generation via NFAT1 activation in Alzheimer’s disease animal model. *FEBS J* 23, 2639–2649 (2009).

31. Sastre, M. et al. Nonsteroidal anti-inflammatory drugs and peroxisome proliferator-activated receptor-gamma agonists modulate immunostimulated processing of amyloid precursor protein through regulation of beta-secretase. *J Neurosci* 23, 9796–9804 (2003).

32. Sastre, M. et al. Nonsteroidal anti-inflammatory drugs repress beta-secretase gene promoter activity by the activation of PPARgama. *Proc Natl Acad Sci U S A* 103, 443–448, doi:10.1073/pnas.0503839103 (2006).

33. Marwarha, G., Raza, S., Prasanthi, J. R. & Ghribi, O. Gadd153 and NF-kappaB regulates translation of JAK2 and ERK1/2 signaling pathways and direct binding of STAT1 to BACE1 promoter in astrocytes. *Glia* 55, 253–262 (2007).

34. Chen, C. H. et al. Nuclear factor-kappaB regulates betaAPP and beta- and gamma-secretase-mediated AD-related BACE1 expression in neurons. *J Neurochem* 98, 85–96, doi:10.1111/j.1471-4159.2006.03692.x (2006).

35. Bourne, K. Z. et al. Nonsteroidal anti-inflammatory drugs repress beta-secretase gene promoter activity by the activation of PPARgamma. *Proc Natl Acad Sci U S A* 103, 443–448, doi:10.1073/pnas.0503839103 (2006).

36. Marwarha, G., Raza, S., Prasanthi, J. R. & Ghribi, O. Gadd153 and NF-kappaB crosstalk regulates 27-hydroxycholesterol-induced increase in BACE1 and beta-amyloid production in human neuroblastoma SH-SY5Y cells. *PLoS One* 8, e70773, doi:10.1371/journal.pone.0070773 (2013).

37. Chen, C. H. et al. Increased NF-kappaB signalling up-regulates BACE1 expression and its therapeutic potential in Alzheimer’s disease. *Int J Neuropsychopharmacol* 15, 77–90, doi:10.1017/S1461145711000149 (2012).

38. Bourne, K. Z. et al. Differential regulation of BACE1 promoter activity by nuclear factor-kappaB in neurons and glia upon exposure to beta-amyloid peptides. *J Neurosci* 35, 1194–1204, doi:10.1523/jneurosci.21252.2007 (2017).

39. Shami, L. et al. Nuclear factor-kappaB regulates betaAPP and beta- and gamma-secretases differently at physiological and supraphysiological Abeta concentrations. *J Biol Chem* 287, 24573–24584, doi:10.1074/jbc.M111.338305 (2012).

40. Kadonaga, J. T. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol* 1, 40–51, doi:10.1002/wdev.21 (2012).

41. Lenhard, B., Sandelin, A. & Carninci, P. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 21, 2339–2422, doi:10.1093/bioinformatics/bti347 (2005).

42. Juven-Gershon, T., Hsu, J. Y., Theisen, J. W. & Kadonaga, J. T. The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol* 20, 253–259 (2008).

43. Yamashita, R., Sugano, S., Suzuki, Y. & Nakai, K. DBTSS: Database of Transcriptional Start Sites progress report in 2012. *Nucleic Acids Res* 40, D150–154, doi:10.1093/nar/gkt1005 (2012).

44. Ge, Y. W., Maloney, B., Sambamurti, K. & Lahiri, D. K. Functional characterization of the 5′ flanking region of the BACE1 gene identification of a 91 bp fragment involved in basal level of BACE promoter expression. *FEBS J* 18, 1037–1039 (2004).

45. Lammich, S., Schobel, S., Zimmer, A. K., Lichtenenthaler, S. F. & Haass, C. Expression of the Alzheimer protease BACE1 is suppressed via its 5′-untranslated region. *EMBO Rep* 5, 620–625 (2004).

46. Rogers, G. W., Jr., Edelman, G. M. & Mauro, V. P. Differential utilization of upstream AUGs in the beta-secretase mRNA suggests that a shunting mechanism regulates translation. *Proc Natl Acad Sci U S A* 101, 2794–2799, doi:10.1073/pnas.0308576101 (2004).

47. Zhou, W. & Song, W. Leaky scanning and reinitiation regulate BACE1 gene expression. *Mol Cell Biol* 26, 3353–3364 (2006).

48. Mihailovich, M., Thermann, R., Grohovaz, F., Hentze, M. W. & Zacchetti, D. Complex translational regulation of BACE1 involves upstream AUGs and stimulatory elements within the 5′ untranslated region. *Nucleic Acids Res* 35, 2975–2985, doi:10.1093/nar/gkm191 (2007).

49. Kim, J. & Kim, J. Y.Y.′s longer DNA-binding motifs. *Genomics* 93, 152–158, doi:10.1016/j.ygeno.2008.09.013 (2009).

50. Maston, G. A., Evans, S. K. & Green, M. R. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet* 7, 29–59, doi:10.1146/annurev.genom.7.080505.115623 (2006).

51. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–461, doi:10.1038/nature13187 (2014).

52. Consortium, F. et al. A promoter-level mammalian expression atlas. *Nature* 507, 462–470, doi:10.1038/nature13182 (2014).

Acknowledgments
This work was supported by the National Basic Research Program (973 Program) of China [2012CB517903 to W.Z.]; the Natural National Science Foundation of China [81070269 to W.Z., 81116120498 to T.L and W.Z.]; and the Canadian Institutes for Health Research (CIHR) (MOP-97825 and TAD-117948 to W.S.).

Author contributions
W.Z. designed the research, performed experiments and wrote the manuscript; W.S. performed the research and wrote the manuscript; Y.X. and S.M. performed experiments; J.W., S.L. and J.L. constructed plasmids; H.L. constructed plasmids and analyzed data; T.L. discussed the project and provided reagents.

Additional information
Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xiang, Y. et al. Two novel DNA motifs are essential for BACE1 gene transcription. *Sci. Rep.* 4, 6864, DOI:10.1038/srep06864 (2014).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/