Gene 501 (2012) 63–78

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Structural and functional characterization of H2 haplotype *MAPT* promoter: Unique neurospecific domains and a hypoxia-inducible element would enhance rationally targeted tauopathy research for Alzheimer's disease

Bryan Maloney [a], Debomoy K. Lahiri [a,b,*]

[a] Laboratory of Molecular Neurogenetics, Department of Psychiatry, Institute of Psychiatric Research Indiana University School of Medicine, Indianapolis, IN 46202, USA
[b] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA

ARTICLE INFO

Article history:
Accepted 20 January 2012
Available online 30 January 2012

Keywords:
Cell type specificity
Gene regulation
H2 haplotype
Neurofibrillary tangles
Promoter
Tau

ABSTRACT

Alzheimer's disease (AD) is the leading cause of dementia in the elderly. Extraneuronal plaque comprising mostly the amyloid β peptide and intraneuronal tangles of hyperphosphorylated microtubule-associated τ protein (τ, gene *MAPT*) are typical of AD. Misfolded τ is also implicated in Parkinson's disease and frontotemporal dementia. We aim to understand the regulation of the human *MAPT* promoter by mapping its functional domains. We subcloned a 4868 base pair (bp) fragment from human BAC RPCI-11 100C5. Sequence analysis revealed an H2 haplotype *MAPT* promoter, 5′-UTR, and intronal fragment. Database analysis of the fragment showed 50%–75% homology with mouse and >90% with rhesus monkey. Comparison with human H1 sequences revealed differences that crossed predicted transcription factor sites. DNA-protein interaction studies by electrophoretic mobility shift assay suggested hypoxia response and an active specificity protein 1 (SP1) site in the 5′-untranslated region. Transfection of a series of *MAPT* promoter deletions revealed unique functional domains. The distal-most had different activities in neuronal vs. non-neuronal cells. We have cloned, sequenced, and functionally characterized a 4868 bp fragment of the human *MAPT* 5′-flanking region, including the core promoter region (−302/+4), neurospecific domains (−4364/−1992 and +293/+504, relative to +1 TSS), and a hypoxia-inducible element (+60/+84). Our work extended functional analysis of the *MAPT* sequence further upstream, and explores cell-type specificity of *MAPT* promoter activity. Finally, we provided direct comparison of likely transcription factor binding sites, which are useful to understand differences between H1/H2 pathogenic associations.

1. Introduction

Alzheimer's disease (AD) is the leading cause of dementia in the elderly (Thies and Bleiler, 2011). AD is characterized by neuropathological traits that include senile plaques consisting primarily of amyloid-β peptide (Aβ) and neurofibrillary tangles consisting primarily of hyperphosphorylated microtubule-associated protein τ (MAP). It has been predicted that over 113 million people will suffer from dementia by 2050 (Brodaty et al., 2011). While the majority of AD research has focused on Aβ, intracellular tangles of microtubule associated protein τ also play an important role in the disorder (De Strooper, 2010). In the present work, we report the structural and functional characteristics of an H2 haplotype τ protein gene (*MAPT*) promoter and the identification of a distal neurospecific domain approximately 2kb upstream of the +1 transcription start site and a proximal hypoxia-inducible element in the 5′-flanking region of the *MAPT* gene.

Tauopathies (τ-related disorders) are not limited to AD, but also include frontotemporal dementia (FTD) (Adamec et al., 2001) and over 20 other disorders (Delacourte, 2008), particularly Parkinson's disease (PD) (Williams-Gray et al., 2009). In AD, τ is hyperphosphorylated, and this hyperphosphorylated τ is the primary component of the τ tangle (Iqbal et al., 2010). Generally, hyperphosphorylated τ dysfunction is relegated to "secondary" status behind Aβ peptide in AD progression, but it has been found that elevation of cerebrospinal τ levels and τ:Aβ ratio correlated with more rapid cognitive decline in AD with mild dementia (Snider et al., 2009). Total neurofibrillary tangles

Abbreviations: AD, Alzheimer's disease; AP, activator protein; *APOE*, apolipoprotein E gene; APP, amyloid β peptide precursor gene; Aβ, amyloid β peptide; bp, base-pair; C/EBP, CCAAT/enhancer binding protein; C6, rat glioblastoma cells; CAT, chloramphenicol acetyltransferase protein; *CAT*, chloramphenicol acetyltransferase gene; CMYC, myelocytomatosis viral oncogene homolog; CREB, cyclic AMP responsive element binding proteins; CRHR1, corticotrophin releasing hormone receptor 1; E2F4, E2F transcription factor 4; EMSA, electrophoretic mobility shift assay or gel shift assay; ER, estrogen receptor; FTD, frontotemporal dementia; HIF, hypoxia-inducible factor; HRE, hypoxia responsive element; *MAPT*, microtubule-associated τ gene; MARE, musculoaponeurotic fibrosarcoma oncogene homolog recognition element; NB, human neuroblastoma SK-N-BE extract or human neuroblastoma SK-N-SH cells; NF-κ-B, nuclear factor κ B; NRL, neural retina leucine zipper; PC12, rat pheochromocytoma/neuronal cells; PD, Parkinson's disease; PDEF, prostate derived Ets transcription factor; PR, progesterone receptor; RAR, retinoic acid receptor; SP, specificity protein; TF, transcription factor; TSS, hnRNA transcription start site; U373, human glioblastoma–astrocytoma; UTR, mRNA untranslated region; τ, microtubule-associated τ protein.

* Corresponding author at: 791 Union Drive, Indianapolis, IN 46202, USA. Tel.: +1 317 274 2706; fax: +1 317 274 1365.
E-mail address: dlahiri@iupui.edu (D.K. Lahiri).

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.01.049

in the CA1 region of the hippocampus correlated better with cognitive decline than did Aβ plaque deposition (Giannakopoulos et al., 2009). Furthermore, loss of medial septum cholinergic neurons in a mouse AD model has been linked to τ pathology, and such loss may mean a critical role for τ pathology in loss of cholinergic neurons typical of the disease (Belarbi et al., 2011). It has been generally presumed that aberrant aggregation is the primary etiological contribution of τ to the tauopathies. However, the possibility may exist that the underlying pathologies could begin due to changes in τ function, with aggregation following (Delacourte, 2008). Significant progress has been made in *pre-mortem* detection and modulation of τ aggregates (Kim et al., 2010), and induction of τ disaggregation may prove a promising therapy (Duff et al., 2010).

While it is a modified version of τ protein, i.e., hyperphosphorylated τ, that is usually associated with neurological disorders, studies have shown that regulatory variations in the τ gene (*MAPT*) and its expression levels may be associated with neuropathology. The C↔G polymorphism at +347, within the intron between the 5′-UTR and exon 1, was associated with upregulation of *MAPT* expression and greater AD risk (Sun and Jia, 2009). In addition to single-nucleotide polymorphisms, *MAPT* has two known large-scale haplotypes, specifically H1 and H2 (Kwok et al., 2004). The haplotype polymorphism has been variously claimed to only exist within *MAPT* (Kwok et al., 2004) or to extend as much as 1.3 megabases (Oliveira et al., 2004). In either case, it is part of a 900 kb chromosomal H1 vs. H2 inversion (Stefansson et al., 2005) that also covers the gene for corticotrophin releasing hormone receptor 1 (*CRHR1*), an important inflammatory pathway gene. H1 consists of several subhaplotypes, all of which are associated with some level of greater risk for PD (Kwok et al., 2004; Seto-Salvia et al., 2011) when compared to H2. H1 drives higher levels of *MAPT* gene transcription in cell culture studies (Kwok et al., 2004). *MAPT* haplotype influence on PD risk is cumulative with genetic and environmental factors such as *APOE* genotype, tobacco smoking, and coffee intake (McCulloch et al., 2008). While specific association between overall onset risk for AD and H1 has been rejected (Seto-Salvia et al., 2011), progression from mild cognitive impairment (MCI) to dementia is faster in patients with both *MAPT* H1 and the *APOE* ε4 allele (Samaranch et al., 2010). On the other hand, H2 is associated with 17q21.31 microdeletion syndrome, one of the more common forms of mental retardation (Rao et al., 2010). Likewise, while H2 does not appear to alter absolute risk for FTD, it does reduce age at onset and accompanies a more severe decline of frontal lobe glucose utilization (Laws et al., 2007) and overall worse FTD prognosis once disease is diagnosed (Borroni et al., 2011). H2 has been claimed to be under positive selective pressure in the Icelandic population (Stefansson et al., 2005).

Early studies of AD molecular biology focused on amino acid-coding sequence variations, such as found in the classic APP_{SWE} (Mullan et al., 1992) and $PSEN_{delE9}$ (Thinakaran et al., 1996) familial AD (FAD) mutations. FAD only represents a small minority of AD cases (Thies and Bleiler, 2011), and age of onset, severity, and rate of progression in the far more common sporadic form of AD have all been linked to levels of apolipoprotein E (*APOE*), *MAPT*, and other AD-associated genes (Sun and Jia, 2009; Yamagata et al., 2001). Specific promoter polymorphisms in the *APOE*, *APP*, β-amyloid site cleaving enzyme 1 (*BACE1*), and *MAPT* genes have been linked to AD risk (Lahiri et al., 2005b; Lv et al., 2008; Maloney et al., 2010; Wang and Jia, 2010; Zuo and Jia, 2009). Therefore, both molecular biology and epidemiology point to the vital importance of investigating promoter regulatory effects in AD etiology. In addition to regulation through the 5′-flanking region, a role for specific species of micro-RNA (miRNA) has been implicated in post-transcriptional regulation of some of these aforementioned genes through the 3′-UTR of their respective mRNAs (Long and Lahiri, 2011, 2012).

In our present work, we have cloned and sequenced a 4868 bp fragment of a human chromosome 17 BAC, containing 4364 bp of the 5′-flanking (promoter) sequence of the *MAPT* gene, the sequence of an exon consisting exclusively of 5′-UTR sequence, conventionally referred to as "exon 0" (Sundar et al., 2007), and a portion of the intron between exons 0 and 1 (GenBank ID: JF412354). Our structural analysis determined that the clone was H2 haplotype. We compared JF412354 with other human *MAPT* 5′-flanking sequences and located both proximal and distal H1 vs. H2 polymorphisms, several of which corresponded to predicted transcription factor (TF) binding sites. A previous direct comparison of H1 and H2 covered 1039 bp, of which 808 bp were considered to be the "promoter" sequence (Kwok et al., 2004). This work was done with a single full-length expression clone constructed. No sequence from that particular analysis has been deposited in available databases.

Our functional study of JF412354 included DNA protein interaction studies using the electrophoretic mobility shift assay (EMSA). For example, double-stranded oligomer probes derived from short sequences within the 5′-UTR were radiolabeled to probe mammalian cell nuclear extracts. One of these probes, corresponding to +60/+84 from the +1 transcription start site (TSS) had specific DNA-protein interaction activity in human neuroblastoma (NB) nuclear extracts. Competition EMSA independently against unlabeled probes for SP1, estrogen receptor (ER), progesterone receptor (PR), and retinoic acid receptor (RAR) eliminated EMSA signal for this probe. Of these TF sites, a GC box/SP site was predicted to exist within +60/+84.

Functional expression studies were carried out by DNA transfection experiments in different cell types. For this, a library of 8 forward and 4 reverse promoter deletion fusion clones was constructed in a promoterless chloramphenicol acetyltransferase (CAT) expression vector. Different recombinant *MAPT* promoter–reporter clones were transiently transfected into human and rat cell lines. All cell transfection results indicated several active sites of significant basal promoter activity. For assays with neuronal-type cultures, the core promoter lies within −302/+4, and extension of the JF412354 sequence upstream of this increased promoter activity until a neurospecific region at −4364/−1992 was included, at which point all neuronal cell cultures had reduction in apparent promoter activity. In addition, the inclusion of "exon 0", which is exclusively in the 5′-UTR, significantly increased apparent promoter activity in NB and PC12 cell cultures. We used this data to determine functional domains within the JF412354 sequence, looking at both k ratio (Bayesian) and pairwise *g* for each comparison. When k ratio was 100 or greater and *g* was $\geq 1.0\ \sqrt{MSE}$ (square root of mean square error for the ANOVA) between adjusted signals for compared clones, the clones were deemed to belong to different "domains".

Finally, we aligned JF412354 vs. H1 and H2 human *MAPT* sequences in GenBank, and against corresponding sequences for chimpanzee, gorilla, orangutan, marmoset, mouse, and rat to estimate a phylogenic tree. Our proposed phylogenic tree agreed with several already published estimations that had set the divergence of human *MAPT* from chimpanzee before any split between H1 and H2 haplotypes within humans. Our data also supports the hypothesis that H1 and H2 lineages arose simultaneously from a third, ancestral sequence.

2. Methods

2.1. Reagents

Unless otherwise specified, reagents were purchased from Sigma (St. Louis, MO) and were of "molecular biology" or "analytic" quality. Restriction enzymes were purchased from New England Biological (Ipswich, MA). Other enzymes were purchased from Roche (Indianapolis, IN). Cell culture reagents were purchased from Invitrogen (Carlsbad, CA).

2.2. Statistical analysis

Data were analyzed with the R environment for statistical computing (R Development Core Team, 2011), using the "nortest"