1. ABSTRACT

Recent advances in proteomic, transcriptomic and genomic technologies have revealed much about the ACT protein and gene. In this review, we summarize our current understanding of the structure and potential physiological roles of the ACT protein, catalogue the regulatory elements that have been implicated in expression of the ACT gene, describe its tissue-specific expression and list the single nucleotide polymorphisms (SNPs) within the gene that track ACT variability. The ACT gene has been implicated in a number of complex human disorders and its potential involvement as a risk factor for Alzheimer’s disease has been the subject of intensive research. However, due to previous limitations in methodologies and inadequate sample numbers the data has been conflicting with many studies failing to be replicated. In this regard, we highlight some potential approaches, which may prove to be beneficial in future studies.

2. INTRODUCTION

The serine protease inhibitor, alpha-1 antichymotrypsin (ACT, SERPINA3; MIM 107280) has been implicated in the pathology of a number of devastating human diseases including chronic obstructive pulmonary disease (COPD), Parkinson’s disease (PD), Alzheimer’s disease (AD), Stroke, Cystic Fibrosis, Cerebral Haemorrhage and Multiple System Atrophy. The majority of research has concentrated on the role of ACT in AD, where ACT has not only been detected in the senile plaques and surrounding astrocytes but has been found to form a complex with toxic beta-amyloid (Abeta) (1-42) peptide. Transgenic mice studies have emphasised the importance of ACT in AD pathology; ACT/APP expressing mice showed increased age-related plaque deposition compared with mice expressing APP alone (1, 2). In addition, a recent microarray study (3) has revealed that ACT treatment in combination with Abeta has a more
Phylogenetic analysis using protein sequences classified the human serpins into nine groups (A-I) (6). ACT protein is one of 13 extracellular proteins in group A and is comprised of 423 amino acids including a 25 residue signal peptide at the amino terminus which is cleaved from the mature protein (7, 8). The total molecular weight of ACT is approximately 55-66kDa due to heavy glycosylation at multiple sites (8). This variable glycosylation is typical of plasma serpins but is not required for serine protease inhibitor activity (8).

ACT has a typical serpin structure, comprising of three beta sheets, 8 alpha helices and an active site situated within a hypervariable reactive centre loop (RCL) (Figure 1A-C; (10, 11)). In human ACT this RCL domain is 23 amino acids long and is located near the carboxyl terminus of the protein. The active site of ACT is situated between the leucine 358 and serine 359 residues (12, 13). The length of the RCL is critical for ACT function; relocation of the active site by a single residue within the RCL has been shown to dramatically affect the serine protease inhibitor activity of ACT (14).

Serpins reside in the relaxed (native) conformation until the targeted protease binds to the RCL, which triggers a conformational change as the RCL inserts into one of the beta sheets (beta sheet A), and then migrates to the opposite pole of the serpin (Figure 1D). This process is irreversible and the serpin is no longer able to bind to additional proteases (15). The resulting complex is more stable with higher resistance to heat and SDS-PAGE than the native form. There are several key domains essential for the ability of serpins to undergo this conformational change, the sequence of which, are highly conserved amongst the inhibitory serpins. The hinge allows rapid insertion of the RCL into the beta sheet and is recognized by a consensus sequence amongst the inhibitory serpins (Figure 1A; (16)). The breach and shutter domains facilitate the opening of the incorporating beta-sheet (17). Crystal structure analysis of a Leu55 mutant within the shutter domain of ACT, led to the identification of the intermediate readily polymerised delta form (Figure 1C) with a partially inserted RCL (18).

ACT binds cognate proteases to form a serpin-protease complex, which is cleared from the circulating plasma by the liver, at a rate 10-50 times more rapidly than ACT alone (19), a process that can be blocked by competition of other serpin-protease complexes (20). In addition, formation of this serpin-protease complex increases the expression of the native serpin (21).

Recently, the complete crystal structure of the native mouse orthologue of ACT (serpina3n) has been elucidated (Figure 1B; (22)). The crystal structure for human native ACT has yet to be been published although there are crystal structures characterised for human ACT in the alternative delta (Figure 1C) and cleaved conformations (18, 23). Comparison of the mouse ACT structure with the native crystal structure of the human archetypal serpin (alpha-1 antitrypsin) (24), shows that there is complete overlay of the two tertiary structures with the exception of...
3.1. The physiological role of ACT protein

ACT acts as an inhibitor of several serine proteases including pancreatic chymotrypsin, leukocyte cathepsin G, mast cell chymases, human glandular kallikrein 2, kallikrein 3 (prostate specific antigen), pancreatic cationic elastase and an uncharacterised lung serum protease (25-29). The strongest association is found with cathepsin G and is thus thought to be its major target (30). Cathepsin G is contained in neutrophil granules and is released at the site of inflammation, where it kills and degrades pathogens, remodels tissues and activates pro-inflammatory cytokines and receptors (31). Excessive or prolonged activity of cathepsin G, caused by insufficient serpin regulation, can lead to tissue damage. ACT is a typical acute phase protein, with the amount of circulating protein dramatically increasing in response to inflammation (32).

In addition, ACT has the unique ability amongst serpins to bind to DNA, a property which is independent of its serine protease inhibitor activity (33). Lysine repeats within close proximity of each other in the crystal structure, (between residues 210-212 and 391-396 of ACT) have been identified as the key residues for DNA binding (33). These residues are not present in other serpins. Specific ACT recognition sequences within DNA have not been identified and any functional significance of ACT binding to DNA has yet to be clarified (33).

4. THE ACT GENE

4.1. ACT gene structure and location

The ACT gene is located on chromosome 14q32.1 within a serpin gene cluster containing ten other serpin genes (Figure 2). ACT resides within a distal sub-cluster of this region containing kallistatin (SERPINA4), proteinase C inhibitor (SERPINA5) and the kallistatin like gene (SERPINA13). All of the genes within this distal sub-cluster, including ACT, are highly expressed in the liver but also have additional tissue-specific patterns of expression. Recently, microarray expression profiles have been generated for ACT (34) and show a wide range of expression in multiple tissue types (Figure 3). Importantly, ACT is the only gene within the distal serpin sub-cluster to be expressed in astrocytes (35).

Matrix attachment regions (MARs) are AT rich, highly repetitive and often intergenic regions that bind to nuclear matrices in vitro and are important for chromatin structure. A Mar-wiz algorithm was used to predict MARs in the distal ACT residing cluster (36). Ten potential MARs were identified in this distal cluster including one approximately 4kb upstream of ACT. DNase1-hypersensitive sites (DHSs) are more accessible to trans-acting factors and act as markers for regulatory regions such as promoters and enhancers. Twelve DHSs were mapped to the ACT distal cluster in HepG2 cells, six of which were present in astrocyte cells and all were absent in non-expressing HeLa cells (35). The six astrocyte-specific DHSs were specifically located 5’ to the ACT gene in the region between the transcription start site and the -13kb enhancer suggesting that in astrocytes the other distal cluster members are packaged in inaccessible chromatin.

The ACT gene exon-intron structure is conserved amongst group A serpins and a full, detailed schematic of the gene can be seen in Figure 4A. The gene, 11.66 Kb in length, is comprised of five exons (the first of which is non-coding) and a 3’ untranslated region within the final fifth exon (37). Human ACT is transcribed from a single promoter and no alternative transcripts have been reported (8).
Figure 3. Gene expression profiles of ACT. A) ACT expression profiles in 12 normal tissues retrieved from Gene Expression Omnibus (www.ncbi.nih.gov); entry GDS425 (34), probe set 75248_at. Tissues were pooled from 11-25 individuals and hybridised to a U95 Affymetrix array in duplicate. The height of the bar represents the log to the base 2 of the signal value. The colour of the bar represents the detection call; red bars have present calls whereas pink bars are absent.

4.2. ACT Gene regulation

ACT gene regulation is principally controlled from the promoter region directly upstream from the transcription start site. The promoter elements of ACT have not yet been fully characterised, although there is a TATA box-like sequence at positions -23 to -29 (Figure 5) and an initiator element-like sequence at positions -2 to +6 (37). However, these sequences remain putative as there is no direct supporting experimental data. Several transcription factors influence ACT expression, including the liver specific transcription factor hepatocyte nuclear factor 4 (HNF4)(37), yet their exact binding sites have yet to be mapped (39).

ACT has two upstream cell specific enhancers located approximately 11.5kb and 13kb upstream from the transcription start site (Figure 5; (40, 41)). Both enhancers contain adjacent activator protein 1 (AP-1) and nuclear factor-1-X (NF1) binding sites, which are critical for basal astrocyte expression of the ACT gene (42).

As an acute phase protein, ACT gene expression is heavily modulated by cytokines. IL-6 increases the expression of ACT in hepatic cells (23, 43), but not in bronchial epithelial cells (44) or astrocytes (45). Although astrocytes do not appear to possess a functional IL-6 receptor, co-treatment of human astrocytes with soluble IL-6 receptor (sIL-6R) resulted in potent stimulation of ACT expression (45). Oncostatin M (OSM) stimulates the production of ACT in astrocytes, an effect that is lost following deletion of two STAT binding sites (mapped to positions –125 to –117 and –95 to –87 upstream of the ACT gene (Figure 5)); suggesting that astrocytes possess a specific OSM receptor and that OSM stimulation is mediated via the JAK-STAT pathway (45).

IL-1 has been shown to increase the expression of ACT in hepatic cells (23, 43), bronchial epithelial cells (44) and astrocytes (45). Functional activity mapping experiments with astrocytes revealed that the IL-1 response elements were not within or immediately 5' to the ACT gene. Instead, the IL-1 response is mediated through the -13kb enhancer upstream from the ACT gene. This enhancer contains two nuclear factor kappaB (NFkappaB) binding sites in addition to an AP-1 site (Figure 5; (40, 45)). Disruption of any of these sites results in a reduced ability to respond to IL-1 and addition of NFkappaB inhibitor (IkappaB) totally halted this IL1 response. Interestingly, the recently discovered -11.5kb enhancer also contains a putative NFkappaB site although this has yet to be shown to be functionally active (Figure 5; (41)).

An evaluation of the effect of cytokines on ACT gene expression has recently been evaluated in our laboratory (unpublished observations). Quantitative PCR was used to construct both dose- and time-response curves measuring the responses of HepG2 cells to the cytokines IL-6, IL-1 and OSM. All three cytokines increased expression of the ACT gene in a dose- and time-dependent manner, although the timing and magnitude of the effect differed between the three cytokines (Figure 6). OSM was a potent stimulator of ACT expression reaching a maximal 200-fold increase in gene expression after 16 hours of stimulation. IL-6 induced a maximal 100-fold effect at the slightly later time point of 24 hours. Although IL-1 stimulation significantly increased ACT expression 7-fold, this effect was considerably less than the IL-6 family members.

4.3. ACT genetic variation

Figure 4B shows the position of all single nucleotide polymorphisms (SNPs) located within the ACT gene and its 13kb 5' flanking region that had been submitted to dbSNP prior to July 2006 (www.ncbi.nlm.nih.gov/projects/SNP). Of these 99 SNPs, 41 have been genotyped by the International HapMap
Figure 4. Scale map of the ACT gene and 13 kb 5' flanking region. A) 5' enhancers (blue), promoter (green), exons (red) and transcription factor binding sites (HapMap database and references (40-42, 45)). B) dbSNPs in HapMap database C) Tag SNPs picked using Tag SNP picker (HapMap). D) Phased haplotypes generated from genotyped SNP data in HapMap (each chromosome of the individuals sampled by the project is represented as a line one pixel high and each SNP allele is arbitrarily coloured blue or yellow). E) LD plot (D') for each pair of genotyped SNPs in the HapMap database.

Consortium, as shown highlighted in blue. Using this genotype information, the HapMap Consortium have made information regarding linkage disequilibrium (LD) between these SNPs available on-line (www.hapmap.org). Using the ‘Tag SNP picker’ utility of the HapMap website, informative ‘tag’ SNPs for any region of interest can be
Figure 5. Location of regulatory elements within the ACT gene. The boxes represent regulatory elements with the corresponding transcription factors listed below. The proposed TATA box is underlined and two mapped polymorphic bases (SNPs) within the region are coloured red.

downloaded. Figure 4C shows the 11 ‘tag’ SNPs selected by HapMap (Phase II, Jul06) covering the ACT gene and 5’ flanking region (Chr14; position 94134265…94160508) with a linkage disequilibrium (LD) $r^2$ cut-off >0.8 and a minor allele frequency (MAF) >5% in the European Caucasian (CEU) population. These 11 “tag” SNPs give >80% coverage of the total common variation within this region and this knowledge will reduce the genotyping costs in future genetic association studies. It is important to note that the ‘Tag SNP picker’ software places all SNPs with LD values of $r^2$ >0.8 and MAF >5% into the same ‘SNP bin’ such that any one of these SNPs could be used as the ‘tag’ for that bin. This means that if a reliable assay cannot be designed for a particular polymorphism, a different SNP from within the same bin can be used instead.

Figure 4D demonstrates the phased haplotypes constructed from the genotyped SNPs within this region in the CEU population, as taken from HapMap. In 60 individuals tested by the International HapMap consortium,
Figure 6. Quantitative PCR measurements of the effects of cytokines A) IL-1, B) OSM and C) IL-6 on ACT expression in HepG2 cells. Time is on the X axis and fold change over basal on the Y axis. The colour of the line indicates the cytokine dose (see key). The error bars represent the standard errors from three experiments conducted in triplicate.
The majority of research into the role of ACT in disease has focused on AD pathology. Alzheimer’s disease, the most frequent cause of dementia, is a neurodegenerative disorder characterised by the formation of senile plaques and neurofibrillary tangles in the brain. Senile plaques result from the extracellular deposition of beta-amyloid protein (Abeta). Intact Abeta is proteolytically cleaved from amyloid precursor protein (APP), by proteases referred to as beta- and gamma-secretases (61) to produce both 40 and 42 amino acid peptides (Abeta(40), Abeta(42)), of which Abeta(42) is the principal component of plaques in human brain (62). The intracellular Abeta(1-42) isoform is a pro-amyloidogenic peptide that either directly or indirectly induces neuronal cell death (63), resulting in severe cognitive impairment at the clinical level.

ACT was first linked to Alzheimer’s disease when it was identified as a major component of amyloid brain deposits (64). This finding was supported by the fact that ACT is highly expressed in regions of the brain associated with AD including the hippocampus (64). Furthermore, in vitro studies have demonstrated that ACT forms a complex with the putative neurototoxic Abeta(1-42) in a dose-dependent manner (58, 65, 66). The amino end of Abeta inserts into one β sheet (sheet C) of ACT and the carboxyl terminal into another (sheet A) (67). This transforms the ACT conformation into a stable form, similar to other serpin-protease complexes, and results in loss of ACT inhibitory activity.

The exact role of ACT in senile plaque pathology is as yet undetermined. Some studies report that ACT concentrations can destabilise fibrils as well as break down pre-formed fibrils (58, 65, 68), thus constituting part of the inflammatory response to Abβ aggregation. Stimulation of neuronal and glial cells with ACT/Abeta mixtures have shown that a specific ACT/Abeta(1-42) complex has multiple cellular effects including the activation of inflammatory transcription factors (66, 69).

A recent microarray study examined the effects of ACT treatment in combination with Abeta(1-42) on gene expression in astrocytes (Figure 7; (3)). Expression profiles were generated from astrocytes exposed to ACT, soluble Abeta(1-42), fibrillar Abeta(1-42), ACT/ soluble Abeta(1-42) and ACT/ fibrillar Abeta(1-42). ACT in combination with Abeta(1-42) had a more profound effect on astrocyte gene expression when compared to the effects of Abeta alone. The soluble and fibrillar molecular forms of the ACT/ Abeta(1-42) combination had opposing effects on gene expression with the majority of genes showing a decrease in expression with the soluble form compared to the majority of genes showing an increase in expression with the fibrillar form. The soluble ACT/ Abeta(1-42) combination decreased the expression of a number of genes involved in neurogenesis and the ubiquitin and proteasome pathways whereas the fibrillar form increased the expression of genes involved in the inflammatory and oxidative stress response. These results suggest that ACT modifies the effects of Abeta on astrocyte gene expression and provides further support for the importance of ACT in AD pathology.
By far the most convincing evidence for a role of ACT in AD has been derived from recent in vivo experiments using transgenic mice expressing astrocytic ACT crossed with APP mutant mice. The resultant ACT/APP double transgenic mice showed an increase in plaque deposition compared with those expressing APP alone inferring that ACT is possibly implicated in Aβ clearance (1). In another ACT/APP mouse model, deposition of Abeta was shown to occur in an age-dependent manner (2) and more recently it has been demonstrated that cognitive impairment in the APP mouse model depends on ACT-catalysed amyloid formation (70). It is possible that different ACT structures may enhance or reduce AD pathology. In additional work, ACT and APOE mice were crossed to APP mice and all the resulting combinations studied (70). The APP/APOE mice had the poorest cognitive ability compared to APP mice with the APP/ACT displaying an intermediary performance. This suggests that ACT has a role in cognitive decline and that APOE and ACT act independently and synergistically to promote amyloid plaque formation. These transgenic mice studies provide support for the theory that ACT is behaving as a molecular chaperone of Abeta and that this may be either by increasing Aβ load or inhibiting Abeta clearance within the cell.

Supporting genetic evidence in humans for a role of ACT in AD pathogenesis has proved to be difficult to obtain and remains controversial. The AlzGene section of the Alzheimer Research Forum website (www.alzforum.org) provides an up-to-date collection of all published AD genetic association studies. Odds ratios for those studies regarding the ACT gene are shown in Figure 8.

In 1995, a common variant (alanine) in the ACT signal sequence (rs4934; base change G (alanine) to A (threonine) at position 2076 relative to the transcription start site) was found to be associated with an increased risk of AD (71). This alanine variant results in a more hydrophobic signal sequence, which is more able to transverse the ER, resulting in increased amounts of mature glycosylated ACT secreted in a transfected rat glioma cell line (72). It has also been reported that alanine homozygotes have a lower age-at-onset of AD (73). However, following these initial observations, a number of studies refuted them (74-78) whilst others were supportive (79-81). One study even reported that the threonine and not the alanine variant was implicated in EOAD (82), increasing the risk of developing AD in APOE ε4 carriers 2-3 fold.

The alanine variant allele (G) at this locus is in strong linkage disequilibrium (D’ 0.97) with the previously reported T-allele of an ACT promoter SNP (rs1884082; base change G-T at position -51), which has been reported to be associated with sporadic and EOAD in both a British and Italian cohort (83). The T-allele of this promoter SNP is associated with a 22% mean higher serum ACT concentration and demonstrates increased functional activity (~30%), as assessed by luciferase reporter gene assays, in hepatocytes and a mixed glial cell population (2). This higher activity T-allele has been associated with cognitive decline in an Italian AD sample collection (83) and recently we have shown that the functional effect of this SNP is more pronounced (T-allele over 200% more activity than the G-allele) in astrocytes (84).

The ACT gene has been screened for additional mutations, and nine SNPs were used for AD association studies and haplotype analysis (8). Two additional polymorphisms at positions 241A and 250T were associated with protective effects of AD. Interestingly, the codon 241A allele was in linkage disequilibrium with the alanine variant of the signal sequence polymorphism. A microsatellite locus in the 5’ flanking sequence has also been shown to be associated with sporadic AD in APOE ε4 carriers (85) but the same cohort failed to show any association with the signal sequence polymorphism.

In addition to AD, ACT gene variation has been implicated in the pathology of other devastating diseases, although findings have also been somewhat conflicting. The majority of studies have concentrated on the signal sequence polymorphism (rs4934). A summary of these findings are shown in Figure 8D. The Ala variant is associated in COPD in a Japanese (86) but not an Italian (87) population. However, both of these studies were small with 55 and 66 COPD cases respectively. Recently, our laboratory performed a large case-control haplotype association study (N=1929), for five SNPs in ACT and reported no association between ACT and COPD (88).

Other neurological illnesses in which the signal sequence SNP has been implicated include Parkinson’s disease (PD) (89), multiple system atrophy (90), stroke (91-93) but these findings have not been replicated in several other studies (81, 94-98).

Other diseases in which negative genetic association studies of the ACT signal sequence polymorphism have been published include leprosy (99),
Figure 8. Odds ratios (OR) and 95% confidence intervals for genetic association studies of the ACT gene in Alzheimer’s disease (adapted from the Alz-gene section of the Alzheimer Research Forum website). A) OR for possession of the Thr variant of the signal sequence polymorphism (rs4934) in AD, a= All AD, b=EOAD, c=LOAD. B) OR for possession of the A10 microsatellite marker variant in AD, a= All AD, b=EOAD, c=LOAD. C) OR for possession of the G allele at the promoter SNP (rs1884082) in AD. D) OR for possession of the Thr variant of signal sequence polymorphism (rs4934) in all other diseases studied. Numbers above each bar represent total number of samples in each study.

Late-onset depression (100), dementia in Down’s syndrome (101), hydrocephalus-related dementia (102) and Creutzfeldt-Jakob disease (CJD) (103).

The difficulty in replication of these ACT genetic associations is a typical characteristic of complex disease genetics and can be due to differences in sample size, method of disease diagnosis, age-at-onset, population and gender. Furthermore, these studies concentrated predominantly on the signal sequence polymorphism within the ACT gene. By only looking at one SNP, additive and/or opposing associations with other SNPs are being ignored.
Such additive or opposing actions may be accounted for by multiple SNP or haplotype analysis. In order to gain enough power to detect the presumably subtle association of any of these individual polymorphisms with disease (e.g. if we assume OR = 1.2-1.4), sample sizes in the 1000s rather than the 100s will undoubtedly be required.

6. PERSPECTIVE

In this review we have catalogued how recent advances in molecular techniques have afforded us a wealth of new information on the structure of the ACT protein and also how the ACT gene is regulated. Currently, we have some understanding of the structure of the ACT molecule and how it compares to other closely related serpin gene family members. Recent HapMap advances have catalogued all of the SNPs within the ACT gene such that we are now able to describe the variation in this region in the human population. Whilst genetic association studies using single SNPs have failed to provide convincing evidence for the involvement of the ACT gene in a number of complex diseases, predominantly AD, it is believed by many that haplotype analysis looking at multiple SNP patterns in large cohorts might be more revealing. For example, it is conceivable that multiple SNPs within any given pathway may influence the pathology of disease. Since ACT has been proposed as a candidate, and we are now able to track its variability, it will be interesting to see if it interacts with any other genes to influence risk in a number of diseases.

7. ACKNOWLEDGEMENTS

Crystal Baker, and Olivia Belbin contributed equally to this work. Crystal Baker is funded by the University of Nottingham and Olivia Belbin is funded by the Alzheimer's Research Trust. Kevin Morgan and Noor Kalsheker would also like to thank the Alzheimer’s Research Trust. Kevin Morgan and Noor Kalsheker would also like to thank the Alzheimer’s Research Trust for Network and Program Grant support. We also thank Prof James Lowe for providing the photograph of astrocytes.

8. REFERENCES

18. Gooptu, B., B. Hazes, W. S. Chang, T. R. Dafforn, R. W. Carrell, R. J. Read & D. A. Lomas: Inactive conformation of the serpin alpha (1)-antichymotrypsin indicates two-stage insertion of the reactive loop:
2833

SERPINA3


**Key Words:** Serpin, alpha-1 Antichymotrypsin, Nervous System, ACT gene, Single nucleotide polymorphism, SNP, Alzheimer’s disease, Review

**Send correspondence to:** Dr Kevin Morgan, The University of Nottingham, Division of Clinical Chemistry, Molecular Medical Sciences, Institute of Genetics, University Hospital, Queen’s Medical Centre, Nottingham, NG7 2UH, UK, Tel: 0115 8230724, Fax: 0115 9709167, E-mail: Kevin.morgan@nottingham.ac.uk

http://www.bioscience.org/current/vol12.htm