A bioinformatics approach to the study of the transcriptional regulation of AMPA Glutamate receptors (GRIAs) and genes whose expression are co-regulated with GRIAs.

Allen K.S. Chong

Thesis presented in fulfilment of the requirements for the Degree of *Doctor Philosophiae* at the South African National Bioinformatics Institute, Faculty of Natural Sciences, University of the Western Cape

April 2009

Advisor: Prof. Vladimir Bajic

KEYWORDS

Neuroscience, bioinformatics, transcription, promoter, transcription factor binding site, composite element, gene expression, AMPA glutamate receptor, neurotransmitter, coexpression.

ABSTRACT

It was postulated that each gene has three main sets of transcriptional elements: one which is gene-specific, one which is family-specific, and a third which is tissue-specific. The starting hypothesis for this project had been: "Each family of genes has a distinct set of transcriptional elements that is unique onto this family". The primary aim of this project was therefore the identification of the family-specific set of transcriptional elements within the AMPA receptor gene family. The question then is how does one measure or identify this uniqueness within the promoters of this family of genes. The answer seemed to lie in making an assessment of the promoters of this family of genes against a background of a comprehensive set of promoter sequences and in the process, to try to find the transcriptional elements that were present in the AMPA receptor gene promoters but were not so common in the general population of gene promoters.

To achieve the primary aim of this project, it was essential that a comprehensive dataset of promoter sequences was available. There are ample data freely available through the web. However, it is often not available in a form that we might want it in. Another problem that one constantly encounters is the lack of general consensus among the research community in agreeing on a standard annotation. For example, a gene can sometimes be given 2 or 3 different names by different laboratories which have successfully cloned the same gene. This, in turn, hinders the data collection process. At the start of this project, there was an existing curated database of experimentally-verified eukaryotic promoter sequences called the Eukaryotic Promoter Database (EPD) and a software called Promoter Extraction from GenBank (PEG) which, as its name implies, extracts promoter sequences available through GenBank (Cavin Périer *et al.*, 1998; Zhang & Zhang, 2001; Praz *et al.*, 2002; Schmid *et al.*, 2004). However, limitations existed in both these resources. For EPD, the number of curated promoter sequences available was low and also, the length of these promoter sequences was short. For PEG, the main limitation was that the extraction from GenBank would result in extraction of sequences of variable lengths. Therefore, the 5'-end Information Extraction (FIE) system was developed for the expressed purpose of collecting promoter sequences

without the limitations of PEG. This software relies on the alignment of multiple mRNA/cDNA sequences that are representative of a gene on the human genomic sequence to determine the transcription start site (TSS) of the gene and thus, with this information, extract the promoter sequence for the gene from the available human genomic sequence. This was the first promoter extraction software to work on this principle (Chong *et al.*, 2002). This method was later supported by experimental work carried out by Coleman and colleagues (2002). Using the FIE2 software (Chong *et al.*, 2003), some 10,000-odd human promoter sequences was extracted, starting at 1500bp uptream and ending at 1000bp downstream of the 5'-most TSS.

Following the collection of the human promoter sequences, the approach developed by Bajic *et al.* (2004) was applied to study the promoters of the AMPA receptor genes. This approach relies on both the MATCH program to map putative transcription factor binding sites (TFBSs) to the promoter sequences and a software developed by Bajic *et al.* (2004) that calculates to the density for each TFBS or composite element. Having calculated the densities for the TFBSs and composite elements for both the target promoters (in this case, the AMPA receptor gene promoters) and the background promoters (the 10,000-odd human promoters), the software then calculates the degree of over-representation of each TFBS and composite element in the target promoters (measured against the background promoters) and then ranks the "singles", "pairs" and "triplets" in the order of their degree of over-representation. Using this method, I identified the top 3 ranked "single", "pair" and "triplet" transcriptional elements found commonly within the AMPA receptor promoters. In addition, a conventional phylogenetic footprinting study was also carried out for the human, mouse and rat GRIA1 promoter to identify key transcriptional elements within this subunit's promoter. While the approach developed by Bajic *et al.* (2004) identifies key family-specific transcriptional elements, the phylogenetic footprinting study helps identify key gene-specific transcriptional elements. Thus, they complement one another.

The approach developed by Bajic *et al.* (2004) yielded an interesting result. It was found that the combination of the top 3 ranked "single", "pair" and "triplet" transcriptional

elements found in the AMPA receptor promoters were also found in 47 other genes. It was postulated that these 47 genes might, in fact, be co-regulated / co-expressed with the GRIAs and thus, explaining the existence of a shared promoter profile with the GRIA promoters. In support of this hypothesis, supporting evidence was found in published literature that 7 of these 47 genes (VAMP4, Rab3B, FKBP8, 3-OST-3_A, CLSTN3, SOCS1 and IκBβ) might indeed be involved in the expression and functioning of the AMPA receptors.

DECLARATION

I declare that "A bioinformatics approach to the study of the transcriptional regulation of AMPA Glutamate receptors (GRIAs) and genes whose expression are co-regulated with GRIAs." is my own work, that it has not been submitted for degree or examination at any other university, and that all the resources I have used or quoted, and all work which was the result of joint effort, have been indicated and acknowledged by complete references.

Allen Chong. April 2009

ACKNOWLEDGEMENT

I wish thank Prof. Vladimir Bajic whose help, encouragement and understanding has made the submission of this thesis possible.

I also wish to thank all my friends and colleagues at the Institute for Infocomm Research (formerly, Kent Ridge Digital Labs), Singapore, for their contribution to the work that makes up this thesis. Words cannot fully express my gratitude to Profs. Limsoon Wong and Vladimir Bajic for giving me my first break into this new and exciting field of biomedical science.

My thanks also to my external collaborators, Profs. Kwok Pui Choi and Ferdinando DiCunto, for their advice, assistance and contribution.

A huge thanks to my parents whose love and support (both spiritually and financially) has been unwavering and unconditional.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

This chapter is split into three main headings: In the first part (Part A), a review of the glutamate neurotransmitter system, in particular, the AMPA glutamate receptor family and its molecular and functional properties, is given. In the second part (Part B), current understanding of transcriptional controls on AMPA glutamate receptor expression is discussed. Finally, a look at the way AMPA receptors are sent to the cell surface for expression at the synapse and its interactions with various cellular components are given in the third part (Part C).

Part A

THE GLUTAMATE NEUROTRANSMITTER SYSTEM – GENERAL ASPECTS

Glutamate (Glu) has long been recognized as a ubiquitous neurotransmitter of the central nervous system (CNS) and represents the predominant excitatory neurotransmitter system in the mammalian brain (Monaghan *et al.*, 1989). The explosion of this knowledge has partially been due to the application of molecular cloning technology to this area of study which led to the cloning of the first functional ionotropic glutamate receptor by Hollmann *et al.* (1989). The longest-known, and best-studied glutamate receptors are the ligand-gated ion channels, called "ionotropic" glutamate receptors, which are permeable to cations. A family of G protein-coupled glutamate receptors called "metabotropic" glutamate receptors has also been identified (Sugiyama *et al.*, 1987). These metabotropic glutamate receptors, unlike the ionotropic receptors, are not ion channels but instead activate biochemical cascades within the cell.

The ionotropic glutamate receptor family has traditionally been split into three broad groups based on their pharmacological and electrophysiological characteristics. These are the α-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) receptors, the kainate (KA) receptors and the N-methyl-D-aspartate (NMDA) receptors. A role for glutamate receptors in synaptic plasticity has been suggested (Schuman & Madison,

1994). Most models of learning and memory (and even neuronal development and circuit reorganization) propose an alteration of the strength of synaptic connections between neurons. Consequently, long-term potentiation (LTP), the long-lasting increase in synaptic transmission that is induced by intense synaptic activity (Bliss & Lomo, 1973), has been offered as a potential physiological mechanism which can provide the plasticity that would be central to both memory storage and brain development. Neurons stimulated with NMDA, resulting in an influx of Ca^{2+} can release a diffusible messenger, nitric oxide (NO) (Garthwaite *et al.*, 1988). NO has been implicated as a "reterograde signal" which is thought to be made by the postsynaptic neuron during LTP and to diffuse back across the synapse to the presynaptic cell, presumably strengthening that synaptic transmission by enhancing neurotransmitter release. It has been demonstrated conclusively that the *diffusible* NO released not only increases the synaptic transmission of the synapse from which it is produced but also that of neighbouring synapses (Schuman & Madison, 1994).

A large body of evidence indicates that glutamate receptors are also involved in a number of neurodegenerative diseases and the damage that occurs after head injuries. It has been known for decades that glutamate is toxic to neurons in culture and *in vivo*, and many experiments implicate the glutamate receptor as a mediator of these toxic effects of glutamate (Choi, 1988; Meldrum & Garthwaithe, 1990; Martin & Beal, 1992; Marcoux *et al.*, 1992).

Glutamate is a major neurotransmitter also within the retina, which is part of the CNS, and is now accepted to be the neurotransmitter released by photoreceptors (reviewed by Massey, 1990). It has long been known that the acidic amino acids - aspartate and glutamate - depolarize horizontal cells (one of several types of second-order neurons within the retina), however the concentrations of aspartate and glutamate needed to depolarize these cells were between 0.5 and 20 mM in the applied Ringer solution (Dowling, 1987). These extremely high concentrations of glutamate were thought to be beyond possible physiological ranges and would probably be neurotoxic. However, it seemed that there exist in synaptosomal rat brain fractions a potent glutamate uptake system which could help keep the brain extracellular glutamate concentration below

levels that would kill neurons (Kanner & Sharon, 1978). Therefore, it seemed likely that a similar system may exist in the retina and in fact, it was later discovered that such a system did exist in the photoreceptors of the goldfish retina (Marc & Lam, 1981).
In many cases, new GluR genes were identified independently and simultaneously in several laboratories and thus were given different names. To avoid inconsistency, gene names consistent with those suggested by the Human Genome Organization (HUGO) will be adopted for this and later chapters. To date, there are some 129 mammalian GluR subunits (ionotropic and metabotropic), GluR-like proteins and GluR interacting proteins that have been cloned. However, only the ionotropic AMPA receptor subfamily will be reviewed.

1.1 AMPA receptors: GRIA1-GRIA4 (also known as GluR1 - GluR4)

1.1.1 Cloning and molecular features

Traditionally, molecular cloning meant purifying the protein concerned to obtain partial protein sequence information, which may then be used to synthesize oligonucleotide probes ultimately employed for screening cDNA libraries. This approach was unsuitable for the initial cloning of functional glutamate receptors due to the fact that high-affinity, high-specificity ligands required for high-affinity glutamate receptor purification were not available. Hollmann *et al.* (1989) resorted to functional expression cloning, a cloning technique first introduced by Masu *et al.* (1987), in isolating the first glutamate receptor cDNA clone. With the sequence information of this first clone (GRIA1), several groups carried out standard homology screening and polymerase chain reaction (PCR)-mediated screening for related sequences and found closely related receptor genes GRIA2, GRIA3, and GRIA4 (Boulter *et al.*, 1990; Keinänen *et al.*, 1990; Nakanishi *et al.*, 1990; Sakimura *et al.*, 1990).

The four receptor subunits GRIA1-GRIA4 are of similar sizes (approximately 900 amino acids with an average molecular weight of about 97,500 daltons) and significant amino

acid sequence identity (68-74%) exist between these receptor subunits. The C-terminal half was extremely conserved, in marked contrast to that of the subunits of other ligand-gated ion channels, such as the nicotinic acetylcholine (nACh) and γ-aminobutyric acid A ($GABA_A$) receptor families, which show high sequence diversity within this region (Hollmann & Heinemann, 1994). However, like other ligand-gated ion channels, GRIA1-GRIA4 contain a hydrophobic domain at its N-terminus that represent the signal peptide required for membrane insertion and the N-terminal region also contains numerous consensus sites for *N*-glycosylation. Therefore, this N-terminal region is expected to be located extracellularly.

Initial hydropathy plot analysis of GRIA1 revealed several regions that are candidates for transmembrane domains (TMDs) (Hollmann *et al.*, 1989). The hydropathy profile of the region between amino acids 455 and 810 resembled that seen in other ligand-gated ion channels: three closely spaced putative TMDs which are separated by ~175 amino acid residues from a fourth putative TMD which is located close to the C-terminus of the protein. Although the presumed TMD IV is very prominent, the assignment of the three TMDs in the 'TMD cluster' region is less obvious. It was, initially, believed that the big 'loop' between the putative TMD III and TMD IV was on the cytoplasmic side, consistent with the presence of several consensus phosphorylation sites for Ca^{2+}-calmodulin-dependent protein kinase type II (CAMKII) and protein kinase C (PKC) in this domain, and that the C-terminal domain downstream of TMD IV would be extracellular (reviewed by Hollman & Heinemann, 1994). Molnar *et al.* (1993), on the other hand, presented immunological evidence that indicated that the C-terminus of GRIA 1 may be intracellular. In addition, Nakanishi *et al.* (1990) found that two small parts of the glutamate receptor subunit resemble parts of a bacterial protein that binds glutamine. The bacterial protein binds glutamine with a clamshell-like structure. Stern-Bach *et al.* (1994) suggested broad similarity between the two halves of the clamshell and two parts of the glutamate receptor subunit, suggesting that the amino acid-binding sites of the two proteins might be similar. In the glutamate receptor subunit, half the clamshell is the big 'loop' between the putative TMD III and TMD IV. Thus, the 'loop' between TMD III and IV would need to be extracellular in order to bind glutamate and following on from that

logic, the glutamate receptor subunit would therefore need to wind through the cell membrane an odd number of times. It was reported that the 'loop' was glycosylated when synthetic glycosylation sites were introduced throughout the glutamate receptor subunit protein (Hollmann *et al.*, 1994). Since only parts of membrane proteins destined to be outside the cell are exposed to enzymes that add glycosyl groups, this meant that the 'loop' must be located extracellularly. Other evidence suggest that the supposed TMD II does not span the membrane: Bennett and Dingledine (1995) showed that both ends of "TMD II" are degraded by proteases targeted to the intracellular side of the membrane. Similar studies by Wo and Oswald (1994, 1995) on the goldfish kainate receptor are in agreement with the view (1) that the previously thought to be large intracellular loop is really extracellular and (2) that the putative second TMD may not be a true transmembrane spanning region. For this reason, TMD II will from this point on be referred to as the "M II domain". If this latest model proves to be correct, the proposed 'superfamily' hypothesis of ligand-gated ion channels (Grenningloh *et al.*, 1987) would be invalid and glutamate receptors would represent the first of a novel class of ligand-gated ion channels in mammalian neurobiology. This is now a widely accepted model for the glutamate receptor subunits. (Figure 1.1)

1.1.2 Mechanisms creating receptor diversity

1.1.2.1 Alternative splicing

Each of the GRIA1- GRIA4 subunits exists in two different forms created by alternative splicing of a 115-base pair (bp) region and encodes 38 amino acid residues within a conserved receptor domain immediately preceding TMD IV (Sommer *et al.*, 1990). The alternate versions (exons) were termed "flip" and "flop" (although flip is not the reverse of flop as the names unfortunately suggest). The sequences of the two alternate exons are very similar and most nucleotide substitutions are silent changes with respect to the protein sequences. The flip and flop versions differ in only a few (9 to 11) amino acids and these substitutions are often conservative. A tetrapeptide is consistently different between the flip and flop versions, giving their signature sequences N-$(X)_{21}$-GGGD (for

flop) and S-$(X)_{21}$-KDSG (for flip), which are invariant in GRIA1- GRIA4 (Figure 1.1). The two splice variants do not confer different pharmacological properties to the receptors but they do differ in the efficacy of L-glutamate (L-Glu) in activating the receptor: L-Glu activated channels four to five times more effectively when interacting with the flip version (Sommer *et al.*, 1990). It was also reported that this domain exchange does not affect the ligand binding properties of the AMPA receptors.

1.1.2.2 Subunit assembly

Coexpression of GRIA1 and GRIA2 consistently resulted in larger whole-cell currents (Keinänen *et al.*, 1990), with L-Glu causing a fast-desensitizing current followed by a steady-state plateau. In receptor channels formed by coexpression of GRIA1 flip and GRIA2 flop, the fast-desensitizing component of L-Glu-evoked currents was large and the sustained current was small whereas receptor channels composed of GRIA1 flop and GRIA2 flip gave a small fast-desensitizing component and a large sustained current. Thus, it was concluded that the large fast-desensitizing component arose from the GRIA1 flip, whereas the small sustained current was derived from GRIA2 flop (Figure 1.2) (Sommer *et al.*, 1990). This also indicated that in assemblies of two receptor subtypes, one partner can be dominant with respect to the fast desensitizing component while the other determines the steady-state component and that such dominance is irrespective of whether the receptor subtypes are of the flip or flop version.

The two flip-flop versions are equally abundant but show different regional distribution in the rat brain, particularly in CA3 pyramidal cells (only flip) and dentate gyrus granule cells (more flop than flip) of the hippocampus (Sommer *et al.*, 1990). Their expression during development is also differentially regulated (Monyer *et al.*, 1991). The observed developmental switch in rat from predominantly flip variants before birth to flip plus flop variants after birth might reflect a need for more efficient flip receptors during synaptogenesis, which later in development are toned down by the addition of flop variants (Monyer *et al.*, 1991).

1.1.2.3 RNA editing

Another mechanism creating receptor diversity involves RNA editing of a single codon in the messenger RNA (Sommer *et al.*, 1991). A glutamine (Q) residue in the putative M II domain is encoded in the genes for GRIA1- GRIA4, but nevertheless all GRIA2 cDNA clones from adult animals actually contain an arginine (R). Editing at this position has not been observed in GRIA1, GRIA3 or GRIA4. Because only one gene exists for each of the four AMPA receptors and no alternate exons are present for M II, this observation is best explained by RNA editing. Editing of GRIA pre-mRNAs requires a double-stranded RNA (dsRNA) structure formed by exonic and intronic sequences and is catalysed by an unknown dsRNA adenosine deaminase. Double-stranded RNA (dsRNA)-specific adenosine deaminase (DRADA) has been implicated as an enzyme responsible for the editing of RNA transcripts encoding glutamate-gated ion channel subunits in brain (Kim & Nishikura, 1993; Dabiri *et al.*, 1996) though another dsRNA adenosine deaminase, RED1, has also been cloned and shown to edit the Q/R site in GRIA2 more efficiently than DRADA (Melcher *et al.*, 1996). DRADA requires a cofactor protein(s) commonly present even in non-neuronal cells and the accuracy and efficiency of this RNA editing system appear to be determined by the quantitative balance between DRADA, cofactor and substrate GRIA2 RNA (Dabiri *et al.*, 1996). The significance of this editing mechanism lies in the fact that it changes an amino acid that crucially affects ionic permeation by determining the rectification properties of the channel and it's divalent cation permeability (Hume *et al.*, 1991; Verdoorn *et al.*, 1991).

Immediately adjacent to the flip/flop cassette in the extracellular 'loop' between TMD III and TMD IV lies asnother site that also undergoes RNA editing. This results in a change from the encoded arginine (R) to glycine (G) (Lomeli *et al.*, 1994). Editing at the R/G site is specific for GRIA2, 3 and 4 and is about 80-90% complete in the adult rat brain. RNA editing and splicing at the flip/flop site are developmentally regulated and are cooperative in controlling desensitization and recovery rates of AMPA receptor responses (Seeburg, 1996).

1.1.3 Functional properties

The GRIA1- GRIA4 subunits may form homomeric receptor ion channels when expressed in *Xenopus* oocytes or cultured mammalian cells. In ligand binding studies with [^{3}H]AMPA, binding was most effectively competed by quisqualate (QA), followed by glutamate, and least effectively by kainate (Keinänen *et al.*, 1990). EC_{50} (the effector concentration for half-maximal response) also indicated the same rank order of potency:

QA> domoate (DOM) ~ AMPA> Glu > KA

(Hollmann *et al.*, 1989; Nakanishi *et al.*, 1990; Sakimura *et al.*, 1990). Although AMPA is the most potent specific agonist, kainate elicits the largest responses (Hollmann & Heinemann, 1994). Consequently, most expression studies use KA as the standard agonist for GRIA1- GRIA4. It is thought that AMPA, glutamate and quisqualate function as partial agonists whereas kainate and domoate are full agonists. Consistent with this, when kainate (100 μM) was applied in the presence of AMPA (5 μM) to human embryonic kidney (HEK) cells coexpressing GRIA1 and GRIA2, the amplitude of the kainate-evoked current was reduced to 40% of control (Keinänen *et al.*, 1990). Kainate- and domoate-induced steady-state currents are larger because these agonists do not desensitize the receptors. Glutamate, quisqualate and AMPA currents are smaller because of fast desensitization. The desensitization time constants for GRIA2 and GRIA4 are ~36 ms and ~8 ms respectively (Burnashev *et al.*, 1992a). After an initial large peak current, desensitization rapidly reduces the current 2.5- to 8-fold to a steady-state level (Verdoorn *et al.*, 1991).

The GRIA2 homomeric receptor channel stands out among the GRIA1- GRIA4 subunits because responses obtained from oocytes injected with GRIA2 RNA were very small. KA-evoked depolarizations in GRIA2-injected oocytes could only be detected in oocytes injected with large amounts of RNA (10 to 25 ng, as opposed to 2 ng of RNA for GRIA1- and GRIA3-injected oocytes) (Boulter *et al.*, 1990). Also, heteromeric receptors containing the GRIA2 subunit gave a linear (or slightly outwardly rectifying) current-voltage (I-V) relation and it seemed that the slope of I-V curves for receptors containing this subunit were also not as steep as the other AMPA receptors which display a strong inward rectification (Boulter *et al.*, 1990). In fact, the I-V plots for receptor combinations containing the GRIA2 subunit were quite similar to that of oocytes injected with hippocampal RNA.

It was reported that notably larger currents are mediated by heteromeric receptors (Keinänen *et al.*, 1990). However, it has been shown that while KA-evoked currents are generally larger (when comparisons are made with the summed currents for the individual subunits), responses to quisqualate, glutamate and AMPA are significantly reduced in oocytes expressing the subunit combinations relative to GRIA1 (Boulter *et al.*, 1990). Both homomeric and heteromeric channels show cooperativity, which suggests a multimeric structure of the native channel (Nakanishi *et al.*, 1990).

The responses of AMPA receptors (homomeric or heteromeric combinations) to either glutamate (300 μM) or kainate (100 μM) can be potentiated by concurrent application of cyclothiazide (100 μM) (Partin *et al.*, 1993). Even homomeric GRIA2, whose responses to kainate is too small to record, show low amplitude agonist-activated currents in the presence of cyclothiazide. Flop splice variants, however, showed much less potentiation of KA and Glu responses by cyclothiazide (Partin *et al.*, 1993). Furthermore, it should also be noted that cyclothiazide, in addition to preventing desensitization of AMPA receptors, also increases the apparent affinity of AMPA receptors for glutamate agonists (Patneau *et al.*, 1993; Yamada & Tang, 1993).

All recombinant AMPA receptor subunits and their heteromeric combinations, except for GRIA2 and combinations containing GRIA2, were partially permeable to various divalent cations, including Ca^{2+} and Mg^{2+} (Hollmann *et al.*, 1991), as demonstrated by I-V curve shifts to positive potentials in high divalent/ low monovalent cation solutions. Ca^{2+} permeability of AMPA receptors was also demonstrated directly with fluorescence-monitoring of Ca^{2+} influx into GRIA3-expressing oocytes upon channel activation in physiological solutions (Keller *et al.*, 1992). The dominance of the GRIA2 subunit in determining permeability to Ca^{2+} and other divalent ions is attributed to the presence of the positively charged arginine (R) in place of a glutamine (Q) residue within the M II domain. A mutant GRIA2 containing Q instead of R in M II is made Ca^{2+}-permeable with a rectifying I-V curve, whereas a mutant GRIA3 (or GRIA1 or GRIA4) containing R instead of Q has extremely low Ca^{2+} permeability and exhibit a linear I-V relation (Hume *et al.*, 1991; Burnashev, *et al.*, 1992b). Although these data seemed to suggest that rectification properties and divalent cation permeability were coupled, further

mutagenesis studies suggested otherwise. If a histidine (H) or an asparagine (N) was introduced at the Q/R site of inwardly rectifying, Ca^{2+}-permeable GRIA1, GRIA3 or GRIA4, Ca^{2+} permeability was maintained but the subunit exhibited a linear I-V relation (Burnashev *et al.*, 1992a; Curutchet *et al.*, 1992; Dingledine *et al.*, 1992). Rectification of receptors lacking GRIA2, in fact, arises from fast, voltage-dependent channel block by intracellular polyamines (Kamboj *et al.*, 1995; Koh *et al.*, 1995).

GRIA2 may serve an important function in regulating the Ca^{2+} permeability of GRIA1-GRIA4 receptors *in vivo*, during development as well as in abnormal conditions such as neurodegenerative diseases. GRIA2 in neocortex, striatum, and cerebellum has a developmental pattern different from that of GRIA1 and GRIA3. GRIA2 levels increase monotonically relative to those of GRIA1 plus GRIA3, presumably leading to the formation of fewer Ca^{2+}-permeable receptors (Pellegrini-Giampietro *et al.*, 1992a).

1.1.4 Regulation of glutamate receptors

There is increasing evidence for regulation of glutamate receptors by protein phosphorylation (Liman *et al.*, 1989; Greengard *et al.*, 1991; Moss *et al.*, 1993; McGlade-McCulloh *et al.*, 1993; Roche *et al.*, 1996). Phosphorylation enhances glutamate receptor activity while dephosphorylation of the receptor protein has the opposite effect. cAMP-dependent protein kinase (PKA) is highly expressed in postsynaptic densities and anchored at that site through binding of its regulatory subunit to an anchoring protein known as the A kinase anchoring protein-79 (AKAP-79) (Coghlan *et al.*, 1995). The activation of PKA by cAMP causes enhancement of AMPA, kainate and NMDA currents in retinal and brain synapses (Liman *et al.*, 1989; Wang *et al.*, 1991; Greengard *et al.*, 1991; Rosenmund *et al.*, 1994; Schmidt *et al.*, 1995; Roche *et al.*, 1996; Raman *et al.*, 1996). Conversely, inhibition of phosphatases enhances the currents stimulated by both AMPA and kainate (Wang *et al.*, 1991; Wyllie & Nicoll, 1994). Greengard *et al.* (1991) observed an enhancement by forskolin in hippocampal whole-cell recordings of AMPA current and also demonstrated by single-channel analysis that PKA increases the opening frequency and mean open time of non-NMDA glutamate receptors. In white perch retinal

horizontal cells, elevation of cAMP or microinjection of PKA enhanced kainate-evoked currents by increasing the frequency of channel opening and channel open time (Liman *et al.*, 1989). In whole-cell recordings of the perch's horizontal cells, fast application of L-glutamate gave rise to fast transient currents with peak values of 200 pA that desensitized within 100 ms (Schmidt *et al.*, 1995). However, if the cells were incubated with dopamine prior to application of L-glutamate, then desensitization was significantly reduced and L-glutamate induced steady-state currents with amplitudes similar to the previously observed transient currents (Schmidt *et al.*, 1995). It is believed that this dopamine-dependent modulation of the horizontal cell's AMPA receptors is mediated by a cAMP-dependent protein phosphorylation. It has also been shown that GRIA1 can be phosphorylated on multiple sites that were located entirely on the C-terminus of the protein (Roche *et al.*, 1996). PKA was shown to specifically phosphorylate serine-845 and PKC phosphorylated serine-831 of GRIA1 transfected in HEK cells and in neurons in culture (Roche *et al.*, 1996). These results are consistent with the latest proposed topology of glutamate receptors which suggest that the C-terminus is found intracelullarly. PKA phosphorylation resulted in a 40% potentiation of the peak current through GRIA1 homomeric channels (Roche *et al.*, 1996). Cotransfection with v-src tyrosine kinase also produced phosphorylation of GRIA1 at a tyrosine (Y) residue (Moss *et al.*, 1993). However, it was believed, then, that tyrosine phosphorylation in GRIA1 occured at residue Y655 in the 'loop' between TMD III and TMD IV. It is possible that there might be possible that there might be a consensus tyrosine phosphorylation site on the C-terminus which, as yet, has not been located.

The amplitude of channel current of the AMPA receptors is strongly enhanced by phosphorylation with CAMKII and PKC (McGlade-McCulloh *et al.*, 1993; Yakel *et al.*, 1995). GRIA1 itself has been specifically shown to be strongly phosphorylated by CAMKII, weaky phosphorylated by PKC and not phosphorylated by PKA *in vitro* (McGlade-McCulloh *et al.* 1993). There is strong evidence that CAMKII activation results in the elevated phosphorylation of AMPA-type glutamate receptors in response to NMDA receptor stimulation in hippocampal cells (Tan *et al.*, 1994). Treatment of hippocampal cells with glutamate, which activated CAMKII, resulted in enhanced ^{32}P-labeling of CAMKII (since CAMKII activation leads to rapid autophosphorylation) and

GRIAs. The enhanced phosphorylation of CAMKII and GRIAs by glutamate required NMDA receptor stimulation since the NMDA receptor antagonist, AP5, blocked the response. It should be noted that activation of AMPA receptors releases the Mg^{2+} channel block on NMDA receptors, allowing activation of the NMDA receptor. Thus, there is a possible "symbiotic" existence of AMPA and NMDA receptors at the postsynaptic terminals.

On a final note, a more novel model of activity-driven regulation of glutamate receptors may occur at the dendrites. It was reported that various RNAs have been identified in dendrites, including mRNAs encoding neurotransmitter receptors (reviewed by Steward, 1997). It is therefore possible that select dendritic mRNAs, docked at a synaptic target site, can be translated selectively upon demand, for example, as a result of transsynaptic activity or the action of trophic factors (Steward, 1997). This would thus allow functional plasticity of the neurocircuitry giving each of the thousands of synaptic connections of a given neuron an ability to be modulated independently. More conventionally, it was shown that NMDA receptor activation in hippocampal CA1 neurons may contribute to enhanced AMPA receptor-mediated transmission observed during LTP and activity-dependent synaptic maturation by inducing rapid delivery of GRIA1 (and perhaps other AMPA receptor subunits) into dendritic spines as well as clusters in dendrites (Shi *et al.*, 1999).

Part B

TRANSCRIPTIONAL REGULATION OF AMPA RECEPTOR EXPRESSION

1.2 Dynamic expression of GRIA subunits

The regulation of AMPA receptor subunit expression is dynamic and is controlled spatially and temporally. For example, in hippocampus, most AMPA receptors are composed of a combination of either GRIA1/GRIA2 subunits or GRIA2/GRIA3 subunits (Wenthold *et al.*, 1996). The role of glutamate receptors in the development of the nervous system is well-established (Molnar *et al.*, 2002). The expression of AMPA receptor subunits were observed to be differentially regulated during development (Pellegrini-Giampietro *et al.*, 1991; Pellegrini-Giampietro *et al.*, 1992a). In fact, it was recently shown that expression of GRIA1flip alone in architecturally mature dendrites is sufficient to initiate a remodeling of the dendritic arbor (Inglis *et al.*, 2002). The repertoire of glutamate receptors expressed by developing motor neurons differs significantly from the glutamate receptor phenotype of mature motor neurons (Kalb *et al.*, 1992; Stegenga & Kalb, 2001). Neonatal motor neurons express very high levels of the GRIA1flip subunit but not adult motor neurons (Jakowec *et al.*, 1995a, b). Inglis and colleagues (2002) were thus able to show that GRIA1flip might be involved in the modeling of dendritic architecture of motor neurons during development and possibly, this mechanism, with the appropriate stimulus, can be engaged in mature motor neurons to modify the existing dendritic architecture. Interestingly, Zamanillo *et al.* (1999) discovered that in GRIA1$^{-/-}$ adult mice, associative LTP was absent in hippocampal CA3 to CA1 synapses.

Expression of AMPA receptors can also be regulated by extraneous conditions, for example, ischemia (Pellegrini-Giampietro *et al.*, 1992b). Following a transient but servere global ischemia insult, it was found that hippocampal CA1 cells showed a decrease in GRIA2 and GRIA3 mRNAs at 24 hours, prior to any loss of neurons (Pellegrini-Giampietro *et al.*, 1992b). More importantly, the relative reduction in GRIA2 mRNA compared with GRIA1 mRNA levels was observed to be significantly larger in

postischemic CA1 pyramidal cells. This switch in AMPA receptor subunits' mRNA expression was not observed in the CA3 or dentate gyrus regions. Because AMPA receptors containing the GRIA2 subunit has low divalent cation permeability, the relative reduction in GRIA2 expression in CA1 cells leads to increased Ca^{2+}-permeability through AMPA receptors in response to endogenous glutamate and may thus contribute to the delayed necrosis of CA1 neurons (Hollmann *et al.*, 1991; Jonas *et al.*, 1994; Gorter *et al.*, 1997). It would also explain why the CA3 and dentate gyrus regions are resistant to ischemic injury.

Conversely, increased GRIA2 subunit levels can be detected in other regions of the hippocampus (eg. dentate gyrus) which are more resistant to ischemic injury (Pollard *et al.*, 1993; Friedman *et al.*, 1994). In these cases, the upregulation of GRIA2 acts in a protective nature.

Changes in AMPA receptor subunit expression can also be observed following chronic administration of such drugs as pyschotropics, pyschostimulants, antidepressants and antipyschotic medications (Ortiz *et al.*, 1995; Skolnick *et al.*, 1996; Brene *et al.*, 1998; Kelz *et al.*, 1999). For example, chronic exposure to cocaine causes the persistent expression of delta FosB (ΔFosB), which in turn alters AMPA receptor subunit expression in the nucleus accumbens (Hope, 1998; Kelz *et al.*, 1999). Expression of ΔFosB significantly increased levels of GRIA2 by more than 50% in the nucleus accumbens while no change were observed in GRIA1 or NMDA receptor subunits levels (Kelz *et al.*, 1999). The increased GRIA2 levels is believed to be responsible for the enhanced behavioural sensitivity to cocaine since an overexpression of the edited GRIA2 subunit in the nucleus accumbens enhanced the rewarding effects of low dose cocaine (Kelz *et al.*, 1999).

It is long believed that long-lasting changes in synaptic function is the cellular basis of learning and memory (Alkon & Nelson 1990; Kandel 1997). The most thoroughly characterized examples of such changes in synaptic function (synaptic plasticity) in the mammalian nervous system are LTP and long-term depression (LTD). A remarkable

feature of LTP and LTD is that a short period of synaptic activity (either high- or low-frequency stimulation) can trigger persistent changes of synaptic transmission lasting at least several hours and often longer. This single property is what led investigators to initially suggest that these forms of plasticity are the cellular correlate of learning (Bliss & Gardner-Medwin 1973; Bliss & Lomo 1973). A fundamental question, however, remained: Is the change in synaptic strength during these forms of plasticity primarily due to a pre- or postsynaptic modification? This question seem to be answered with the identification of postsynaptically "silent synapses" and the demonstration that they could be converted to active synapses by a postsynaptic modification (Liao *et al.*, 1995; Isaac et *et al.*, 1995; Durand *et al.*, 1996; Liao *et al.*, 1999). Dendrites bearing postsynaptic NMDA glutamate receptors (GRINs), although making a significant number of synaptic contacts with the axonal presynaptic membrane, are said to be postsynaptically “silent” at resting potential because of the voltage-dependent blockade of GRINs by magnesium. However, after an LTP-inducing protocol lasting only minutes, AMPA receptors are expressed and branches that acquire GRIAs are stabilized while those that do not are retracted (Nowak *et al.*, 1984; Mayer *et al.*, 1984; Isaac *et al.*, 1995; Liao *et al.*, 1995; Liao *et al.*, 2001). Thus, wholesale appearance of an AMPA response at such synapses during LTP, with no change in the NMDA response, strongly supports a postsynaptic modification consisting of a functional recruitment of AMPA receptors. However, Hohnke *et al.* (2000) found that although a small number of silent retinogeniculate synapses are present, there is no overall change in GRIA/ GRIN contribution when the retinogeniculate axons from ON-center and OFF-center retinal ganglion cells segregate to form ON/OFF sublaminae in the lateral geniculate nucleus. Thus, they argue against the idea that the conversion of silent to functional synapses could play a role in the development and refinement of inputs. But it would be naive to believe that one mechanism could serve all cases of activity-dependent plasticity. Moreover, what is overlooked and the conclusion that may be drawn from all these studies is the fact that although axon guidance and development may not be dependent on GRIA, but perhaps GRIA-dependent mechanisms may play a role in dendritogenesis.

Thus, we can clearly see here that AMPA receptor expression is a dynamic process.

1.3 The eukaryotic transcriptional machinery

1.3.1 The core promoter and basal transcription factors

Control of gene expression begins at the promoter which is made up by genomic DNA sequences found upstream of the transcription start site (TSS), although it can often include sequences as far off as the first intron. Transcription factors (TFs) recognize short DNA sequence motifs on the promoter called the transcription factor binding sites (TFBSs) and act in concert with one another to either initiate or repress a gene's expression.

Generally, a TF does not just activate the expression of a single gene, but numerous genes. For example, a growth factor usually activates transcription of a group of early genes coding for proteins required for the start of DNA synthesis and cell proliferation. After the completion of transcription, the primary transcript is processed by attachment of a 5'-cap and a 3'-polyadenylyl tail, and splicing to remove the intronic sequences (Stryer, 1995). The processed mature mRNA is then transported from the nucleus to the cytoplasm where it is translated into protein.

There are 3 types of RNA polymerases found in the eukaryotic cells. The enzyme catalyzing eukaryotic mRNA synthesis is RNA Polymerase II (RNA Pol II) (Cramer, 2004; Weil *et al.*, 1979). Yeast RNA Pol II is composed of 12 subunits encoded by the RPB1 to RBP12 genes (Cramer, 2002; Hampsey, 1998). There is extensive structural conservation among subunits of eukaryotic RNA Pol II. In fact, 6 subunits of human RNA Pol II can functionally replace their homologs in yeast (McKune *et al.*, 1995). The two largest subunits, RBP1 (~200 kDa) and RBP2 (~150 kDa), are the most highly conserved subunits.

Following the discovery of RNA Pol II, the general transcription factors (GTFs), defined by being the minimal complement of factors required for reconstituting accurate

transcription from a minimal promoter by RNA Pol II *in vitro*, were also identified (Orphanides *et al.*, 1996). Many different factors have also been identified as transcriptional "co-activators", however, most of these are not general factors required for the expression of all RNA Pol II genes. The most universal cofactor that serves to transduce information between gene-specific transcription factors and the core RNA Pol II machinery is a large, modular complex called the Mediator (Myers & Kornberg, 2000).

Transcription of protein-encoding genes first requires the assembly of a preinitiation complex (PIC) which are made up by 5 GTFs - TFIIB, TFIID, TFIIE, TFIIF and TFIIH - and a sixth, TFIIA, potentiates the magnitude of the transcription (Weinzierl, 1999). The assembly starts with the binding of TFIID to a short AT-rich sequence in the promoter ~30 bp upstream of the TSS called the TATA box (Patikoglou *et al.*, 1999). TFIID is a multisubunit complex comprising of the TATA box-binding protein (TBP) and at least 14 TBP-associated factors (TAFs) (Albright & Tjian, 2000; Green, 2000). The core domain of TBP binds the minor groove of an 8 bp TATA element, unwinding about a third of a helical turn and bending the DNA ~80 Å toward the major groove (Kim *et al.*, 1993a; Kim *et al.*, 1993b). In addition, TAFs may also bind nearby promoter elements such as the initiator element (Inr) and downstream promoter element (DPE). RNA pol II itself recognizes features of the Inr which might assist the correct positioning of the polymerase on the promoter (Carcamo *et al.*, 1991; Weis and Reinberg, 1997). However, *in vitro* transcription and DNA binding experiments using recombinant partial TBP–TAF complexes, revealed that together, TAF1 (also known as $TAF_{II}250$) and TAF2 (also known as $TAF_{II}150$), can mediate core promoter discrimination (Verrijzer *et al.*, 1994, 1995). Individually, neither TAF1 nor TAF2 singles out a clear consensus sequence. However, a combination of these 2 TAFs (TAF1-TAF2) specifically binds the Inr (Chalkley & Verrijzer, 1999). Additionally, UV crosslinking shows that TAF1 and TAF2 are normally positioned close to the Inr, while TAF6 and TAF9 lie close to the DPE (Oelgeschlager *et al.*, 1996; Burke & Kadonaga, 1997).

Binding of TFIID to the promoter is followed by TFIIB's entry which stabilizes TFIID at the promoter by binding TBP via its conserved C-terminal core domain and sequences

flanking the TATA box (Nikolov *et al.*, 1995; Tsai & Sigler, 2000). A subset of eukaryotic promoters contains a TFIIB recognition element (BRE) located just upstream of the TATA box and can stabilize the interaction between TBP and TFIIB onto DNA (Lagrange *et al.*, 1998; Qureshi & Jackson, 1998; Wolner & Gralla, 2001). In fact, Fairley and colleagues (2002) shows that human TFIIB undergoes a conformational change when assembled into a TBP-TFIIB-DNA complex, adopting one conformation for promoters with a BRE consensus sequence and another for promoters lacking this element. In archaea, the BRE is the primary determinant of transcription orientation (Bell *et al.*, 1999; Littlefield *et al.*, 1999).

Like TFIIB, TFIIA also recognizes the TBP-DNA complex (Geiger *et al.*, 1996; Tan *et al.*, 1996). TFIIA also helps stabilize TBP-DNA binding and strongly promotes binding of TFIID to DNA by competing with the TAF1 N-terminal domain that occludes the DNA-binding surface of TBP when TFIID is not bound to DNA (Weideman *et al.*, 1997; Kokubo *et al.*, 1998; Liu *et al.*, 1998). In fact, a considerable change in DNA-binding activity of TFIID is observed in the presence of TFIIA and transcriptional activators (Chi *et al.*, 1995).

The N-terminus of TFIIB contains a zinc ribbon motif that binds to RNA Pol II and thereby recruits RNA Pol II/TFIIF into the PIC (Ha *et al.*, 1993; Pardee *et al.*, 1998). However, transcription still cannot occur until TFIIE and TFIIH are incorporated into the PIC. TFIIH is the largest GTF consisting of nine subunits with well defined enzymatic activities which include DNA-dependent ATPase, ATP-dependent DNA helicase and cyclin-dependent protein kinase (Feaver *et al.*, 1991; Lu *et al.*, 1992; Schaeffer *et al.*, 1993; Serizawa *et al.*, 1993; Roy *et al.*, 1994). The helicase catalyzes the ATP-dependent promoter melting which results in a conformational change that physically separates the two DNA strands to yield an open promoter complex (Douziech *et al.*, 2000; Kim *et al.*, 2000). After promoter melting and transcription initiation, the C-terminal domain of the largest RNA Pol II subunit, RBP1, is phosphorylated by TFIIH's kinase, an event that facilitates promoter clearance and progression into the elongation phase of transcription (Valay *et al.*, 1995).

Most promoters contain one or more of the elements decribed above (e.g., TATA-box), but no one element is absolutely essential for promoter function.

1.3.2 Gene-specific transcription factors

The ultimate target of many signal transduction cascades is the activation of TFs that bind their respective TFBSs in the promoter. A well-studied example is the JAK/STAT pathway utilized by chemokine G-protein-coupled receptors (Mellado *et al.*, 1998; Rodriguez-Frade *et al.*, 1999a, b; Vila-Coro *et al.*, 1999; Soriano *et al.*, 2003). The promoters of eukaryotic genes contain multiple TFBSs, allowing each gene to respond to multiple signaling pathways and facilitating the fine-tuning of transcript levels. The activities of many TFs are also modulated by other TFs bound nearby (Lefstin & Yamamoto, 1998; McKenna & O'Malley, 2002). Thus, a single activated TF can induce transcription of one gene while repressing that of another. This approach where the combination and context in which TFs are present on a gene's promoter defines whether a gene is expressed or repressed allows the cell to respond to a variety of stimuli using the same TF. Transcriptional control in prokaryotes are far simpler since metabolically related genes are clustered and coregulated in common transcriptional units (operons) by a single transcriptional activator or repressor.

The TFBSs are generally 10 to 30 bp long with a small core of nucleotides within the DNA sequence establishing the criteria for a binding site. The sequence outside this core sequence is usually unconserved and because of this inherent variability, TFBSs cannot be efficiently described by their individual sequence. Thus, two common methods are frequently used to describe a binding site for a TF. First, one can use an IUPAC consensus sequence that employs the use of ambiguous symbols (e.g., B to represent G, C, or T; R to represent A or G) to denote the variability of the nucleotide found at a particular position within the consensus sequence. Second, a position weight matix can be used to describe a TFBS. A position weight matrix is represented by a two-dimensional table with one axis representing the relative position within the TFBS sequence and the

other axis representing the 4 different nucleotides (A, G, C and T). For each position within the sequence, a weight (number) is given for each of the four nucleotides to reflect the preference for that nucleotide at that particular position within the TFBS sequence. A position weight matrix is ideally obtained from a set of functionally characterized binding sites for a given TF (Chen *et al.*, 1995; Quandt *et al.*, 1995; Heinemeyer *et al.*, 1999). Binding affinity (and thus, biological significance) is estimated to occur above a certain threshold score. A large collection functional TFBSs with their respective position weight matrices were derived from literature and can be found in the TRANSFAC database (Knuppel *et al.*, 1994; Matys *et al.*, 2003). To date, the number of TFBSs annotated in TRANSFAC stands at 6627 (release 6.0) (http://www.gene-regulation.com/). A number of tools, such as MATCH and MatInspector, make use of the position weight matrices in TRANSFAC to detect putative TFBSs within promoter sequences (Quandt *et al.*, 1995; Kel *et al.*, 2003).

As mentioned above, the activities of many TFs are also modulated by other TFs bound nearby. This has led to the concept of "composite elements" or "promoter modules" (Diamond *et al.*, 1990; Kel *et al.*, 1995; Firulli & Olson, 1997; Kel *et al.*, 1999; Kel-Margoulis *et al.*, 2000; Boehlk *et al.*, 2000; Kel-Margoulis *et al.*, 2002). A composite element is a set of TFBSs found in combination on the promoter and usually, in close proximity to each other that works synergistically or antagonistically to control the expression of a gene. An example of this is the IL-4-responsive element in the SOCS-1 promoter which contains three STAT6 and one Ets consensus binding sequences (Travagli *et al.*, 2004). Ets-1 is confirmed to physically interact with STAT6 and IL-4 responsiveness was either partially or totally abolished following specific mutations. Furthermore, exogenous expression of Ets-1 in conjunction with STAT6 activation strongly inhibited expression of a SOCS-1 promoter-luciferase reporter (Travagli *et al.*, 2004).

Attempts to record these composite elements systematically was first made by the the COMPEL database and is now succeeded by the TRANSCompel database (Kel *et al.*, 1995; Kel-Margoulis *et al.*, 2000, 2002). The collection of composite elements in these

databases are manually curated and are based on proven composite elements described in the literature. This is a slow and arduous task and at last count (Release 6.0 of TRANSCompel), there were only 256 composite elements annotated by TRANSCompel. Obviously, if one thinks of all possible combination of only pairs that can be made from the 4219 TFs currently recorded by TRANSFAC, one would quickly realize that even if only a small fraction of combinations (from the close to 18 x 10^6 possible pairs that can mathematically exist between the 4219 TFs) exist in vivo, the composite elements so far reported in the literature may literally represent a drop in the ocean.

1.4 Promoter elements controlling AMPA receptor expression

To date, little work has been carried out to functionally characterize the promoters of the AMPA receptor subunits. So far, only one paper details the cloning of the rat GRIA1 promoter while 3 other papers studied in detail the regulatory elements within the mouse and rat GRIA2 promoters (Borges & Dingledine, 2001; Brené *et al.*, 2000; Myers *et al.*, 1998; Köhler *et al.*, 1994). Neither GRIA3 nor GRIA4 promoters have been studied.

1.4.1 The GRIA1 promoter

Five kilobase pairs of the rat GRIA1 promoter was cloned and functionally analysed (Borges & Dingledine, 2001) (see Figure 1.3). At least five transcriptional start sites (TSSs) were identified by primer extension and RNase protection assays at -295, -266, -219, -214 and -202 (relative to the first translational initiation site [TIS]). **In keeping with the paper by Borges & Dingledine (2001), all coordinates discussed henceforth in this section 1.4.1 are given relative to the first TIS of GRIA1**. Two other possible TSSs were also reported at -394 and -333, flanking a 64 bp GA repeat, although these do not seem to be as significant since the bands found by both methods were faint (Borges & Dingledine, 2001). The GRIA1 promoter was determined to be mostly neuron-specific since all GRIA1 promoter constructs that were examined showed higher activity in forebrain compared with glial cultures. Furthermore, results with promoter constructs

created either by 5' or 3' deletion indicate that no single region dominates or is essential for GRIA1 promoter activity in neurons. Even small GRIA1 promoter fragments close to the TSSs retain substantial neuronal selectivity, including the shortest constructs, -258/+7 and -209/+8. This is also observed with other promoters of neuronal genes, such as GRIA2, rat β_2-nicotinic acetylcholine receptor subunit, the mouse neural adhesion molecule polysialic acid synthase and rat synapsin II (Myers *et al.*, 1998; Yoshida *et al.*, 1996; Bessis *et al.*, 1995; Chin *et al.*, 1994). Also, the neuronal to glial expression ratio of the most neuron-specific GRIA1 construct was higher than that of GRIA2 due to the low GRIA1 promoter activity in glia compared to GRIA2.

Deletion of the sequences -1395 to -743 or -258 to +8 lowered the neuronal to glial expression ratio by reducing promoter activity in neurons, suggesting that these two regions are neuron-specific regions which help increase expression within neurons. Furthermore, the activity of the neuron-specific region -258 to +8 was found to be orientation-dependent since inverting this region reduced promoter activity in neurons. In contrast, deletion of the region -689 to -459 reduced the neuronal to glial expression ration by increasing activity specifically in glia, suggesting that this is a glial silencing region. In addition, shortening of the 64 bp GA repeat aversely affected expression in glial cultures but did not affect GRIA1 expression in neuronal cultures, however, completed deletion of the GA repeat affected expression in both neuronal and glial cultures with a 55% and 70% reduction, respectively.

The deletion of a 57 bp region (between -743 and -686) containing an N box (CACNAG) saw a significant increase in activity in both neurons and glia which suggested that perhaps the N box reduces expression of GRIA1. GRIA1 is a TATA-less promoter and like most other TATA-less promoters including those of GRIA2, the NMDA ionotropic glutamate receptor subunits, GRIN1, GRIN2B, and GRIN2C and the kainate ionotropic glutamate receptor subunit, GRIK5 (also known as KA2), Sp1 binding sites were found close to the TSS (at -296 and -275) (Miyatake *et al.*, 2002; Borges & Dingledine, 2001; Chew *et al.*, 2001; Brené *et al.*, 2000; Pieri *et al.*, 1999; Myers *et al.*, 1998; Klein *et al.*, 1998; Bai & Kusiak, 1995).

1.4.2 The GRIA2 promoter

Both the mouse and rat GRIA2 promoters has been cloned and characterized (Brené *et al.*, 2000; Myers *et al.*, 1998; Köhler *et al.*, 1994). The rat and mouse GRIA2 promoters show considerable homology (Brené *et al.*, 2000; Myers *et al.*, 1998). In both cases, it was found that the GRIA2 gene could be transcribed from multiple TSSs that were located approximately 300 to 400 nt. upstream of translation ATG codon (Brené *et al.*, 2000; Myers *et al.*, 1998). Furthermore, the 5'-most dominant TSS of the rat GRIA2 gene is in good agreement with the reported 5'-most TSS for the mouse GRIA2 gene (Myers *et al.*, 1998; Köhler *et al.*, 1994). Figure 1.3 gives a schematic diagram of the rat GRIA2 promoter.

Like the GRIA1 promoter, the GRIA2 promoter lacks identifiable TATA and CCAAT boxes (Brené *et al.*, 2000; Köhler *et al.*, 1994). Also like GRIA1's promoter, the GRIA2 promoter is neuron-specific with even the minimal promoter construct (281 bp) showing preferential expression in neuronal rather than glial cells (Myers *et al.*, 1998). In rat, a region of high GC content spans ~150 bp adjacent to the 5'-most dominant TSS and contains consensus recognition sequences for Sp1/Krox-24 and the nuclear respiratory factor-1 (NRF-1) transcription factors (Myers *et al.*, 1998). Studies with both the rat and mouse GRIA2 promoters indicate that the requisite for the minimal promoter is the proximal region upstream of the TSS containing the putative Sp1/Krox-24 and NRF-1 elements (Brené *et al.*, 2000; Myers *et al.*, 1998). Deletion of either of these two elements resulted in a significant drop (on average, 40%) in GRIA2 promoter activity.
RE1-silencing transcription factor (REST), also known as neuron-restrictive silencer factor (NRSF) or X2 box repressor (XBR), is a zinc finger transcription factor which binds the RE1/ neuron-restrictive silencer element (NRSE) and blocks a gene's expression in nonneuronal cells (Shimojo & Hersh, 2004; Scholl *et al.*, 1996; Schoenherr & Anderson, 1995; Chong *et al.*, 1995). However, REST expression is also observed in neurons suggesting that they also function in regulating neuronal gene expression. A NRSE-like element with a 71% identity to the rat SCG10 gene NRSE is found proximal

to the Sp1/Krox-24 and NRF-1 elements in the rat GRIA2 promoter, at between –174 to –194 relative to the 5'-most TSS (Myers *et al.*, 1998; Huang *et al.*, 1999). A construct which had deleted this sequence entirely showed a significant increase in expression in glial cells but not in neurons (Myers *et al.*, 1998). Moreover, normalized results of REST coexpression in cultured neurons showed that GRIA2 promoter activity was significantly reduced in these cells.

Both glial-cell line derived neurotrophic factor (GDNF) and brain-derived neurotrophic factor (BDNF) potently induced murine GRIA2 promoter activity in differentiated SH-SY5Y cells (Brené *et al.*, 2000). However, an overexpression of REST blocked the ability of GDNF to induce GRIA2 promoter activity, without affecting basal promoter activity in the absence of GDNF (Brené *et al.*, 2000). Paradoxically, mutation of the highly conserved GG residues within the NRSE to TT or deletion of the entire silencer element either attenuated or abolished GDNF- / BDNF-induced promoter activity.

Two AP-1 sites plus twelve additional partial AP-1 sites were predicted computationally and functional AP-1 sites are believed to exist within the GRIA2 promoter based on 1) studies showing expression of ΔFosB in bitransgenic mice significantly levels of GRIA2; and, 2) electrophoretic mobility shift assays (EMSAs) showing increased binding to GRIA2 promoter fragment, which contains AP-1, upon ΔFosB expression (Brené *et al.*, 2000; Kelz *et al.*, 1999).

Part C

TRAFFICKING OF AMPA RECEPTORS TO THE SYNAPSE

1.5 Delivery of AMPA receptor subunits to the synapse

1.5.1 The exocytic pathway

The secretory pathway compartments can be subdivided into 2 central membrane populations, the endoplasmic reticulum (ER)-Golgi system and the *trans*-Golgi network system (Gleeson *et al*., 2004; Traub & Kornfeld, 1997). The ER-Golgi system performs the folding, oligomerization, and co- and post-translational modifications of proteins transiting the secretory pathway. The ultimate subcompartment of the Golgi complex is the *trans*-Golgi network (TGN). Although the TGN houses enzymes for terminal processing of newly synthesized proteins, its main function is to sort and coordinate protein, lipid and membrane traffic within the secretory pathway. The TGN gives rise to a multitude of membrane carriers for anterograde and retrograde transport of newly synthesized cargo proteins heading to the plasma membrane or to other intracellular organelles and it also receives a steady volume of incoming traffic from endocytic and recylcing pathways. The conventional view is that two routes of traffic emerges at the TGN. One route, the constitutive pathway, delivers proteins to the cell surface while a second, selective pathway sorts protein traffic into the intracellular endosomal membrane system.

N-ethylmaleimide-sensitive fusion protein (NSF) is an ATPase required for vesicular transport throughout the exocytic and endocytic pathways (Morgan & Burgoyne, 1995). NSF's binding to membranes is mediated by the soluble NSF attachment protein (α-SNAP). Using an immunoprecipitation approach, the synaptic SNAP receptors (more commonly known as SNAREs) were identified as vesicle-associated membrane protein (VAMP)/synaptobrevin, syntaxin and synaptosomal-associated protein of 25 kDa (SNAP-25) from the bovine brain (Sollner *et al.*, 1993). SNAREs the most intensely studied vesicle and membrane fusion proteins with well over 30 family members (Sollner

et al., 1993; Morgan & Burgoyne, 1995; Hay & Scheller, 1997; Lledo *et al.*, 1998; Mochida, 2000; Zinsmaier & Bronk, 2001; Chen & Scheller, 2001; Gerst, 2003; Matos *et al.*, 2003). Crystallization of the neuronal SNARE core complex revealed a four-helix bundle structure, with one coil of syntaxin and VAMP, and two coils of SNAP-25 intertwined to form a four-stranded coiled-coil structure (Sutton *et al.*, 1998). This confirms several structural studies that predicted a parallel arrangement of the core complex which supports the hypothesis that formation of the SNARE complex fuses two membranes by bringing them into close apposition (Hanson *et al.*, 1997; Lin & Scheller, 1997; Poirier *et al.*, 1998). To disassemble such a stable complex, ATP is needed to dissociate it into monomeric components. Disassembly is carried out by the ATPase NSF and its adaptor protein, α-SNAP (Hanson *et al.*, 1997; Lenzen *et al.*, 1998; Yu *et al.*, 1998, Yu *et al.*, 1999; Rice & Brünger 1999).

1.5.2 Evidence for GRIA receptor delivery by exocytic pathways.

Two lines of evidence provide support for postsynaptic exocytosis playing a role in AMPA receptor delivery. Firstly, the recruitment of AMPA receptors to the synapse that occurs, for example, in activity-dependent synaptic plasticity at silent synapses, is believed to be driven by calmodulin-dependent protein kinase II (CAMKII) (Isaac *et al.*, 1995; Liao *et al.*, 1995; Durand *et al.*, 1996; Liao *et al.*, 1999; Hayashi *et al.,* 2000; Liao *et al.,* 2001, Shi *et al.*, 2001). Long-term potentiation (LTP) or increase CAMKII is shown to induce delivery of AMPA receptors into synapses of rat hippocampal neurons (Hayashi *et al.*, 2000). And, coincidentally, CAMKII is both necessary and sufficient to generate calcium-evoked dendritic exocytosis (Maletic-Savatic *et al.*, 1998).

Brief activation of NMDA receptors in hippocampal slices can produce a long-lasting (>3 hours) increase in synaptic efficacy, that is, an NMDA-induced LTP (Broutman & Baudry, 2001). In this case, NMDA-induced LTP saw a rapid upregulation of GRIA1 and GRIA2/3 subunits in synaptic membranes. However, both Brefeldin A, an inhibitor of protein trafficking between the Golgi apparatus and cell membranes, and KN-62, a CAMKII inhibitor, completely inhibited the NMDA-induced upregulation of GRIA1 and

GRIA2/3 subunits in synaptic membranes and also, NMDA-induced LTP. The involvement of CAMKII in LTP and dendritic exocytosis therefore support to the idea that the delivery of AMPA receptors to the synaptic membrane might be mediated by an exocytic pathway.

Secondly, a study in hippocampal slices showed that loading postsynaptic cells with toxins that specifically perturb membrane fusion could block LTP (Lledo et al. 1998). Botulinum toxin, which disrupts the membrane fusion machinery by proteolytically cleaving SNAP-25, can greatly reduce the magnitude of LTP (Lledo *et al.*, 1998). Another player in the membrane fusion machinery, α-SNAP, which has been shown to enhance neurotransmitter release in the squid giant synapse, has also been shown to enhance synaptic strength (DeBello *et al.*, 1995; Lledo *et al.*, 1998). In fact, on pathways in which LTP has been saturated, treatment with α-SNAP elicited only a slight increase in synaptic strength (22 ± 12%) as compared with a control naïve pathway (58 ± 11%) (Lledo *et al.*, 1998). Thus, we may infer that LTP occurs through the upregulation of AMPA receptors by a postsynaptic-regulated exocytic pathway employing a mechanism similar to that for used for the release of neurotransmitters. Passafaro and colleagues (2001) further proposed that GRIA1 controls the exocytosis while GRIA2/3, the recycling and endocytosis of AMPA receptors.

CAMKII has also been implicated in the regulation of presynaptic exocytosis, acting on actin binding of presynaptic vesicles (Ceccaldi *et al.*, 1995). CAMKII injected presynaptically in squid giant synapse facilitated transmitter release (Llinás *et al.*, 1991). In addition to their well-established postsynaptic action, AMPA receptors also mediate presynaptic effects (Nicoll *et al.*, 2000). Presynaptic AMPA receptors have been shown to modulate synaptic transmission, by depressing the release of inhibitory GABA transmitters in the adult cerebellum (Satake *et al.*, 2000). Schenk and co-workers (2003) studied the delivery of AMPA receptors to the presynaptic membrane of axonal growth cones in hippocampal neurons and demonstrated that, at steady state, a major pool of GRIA1 and GRIA2 subunits is associated with synaptic vesicle membranes. Schenk and

co-workers (2003) provide several lines of evidence to support the idea that insertion of AMPA receptors into presynaptic membranes occurs through an exocytic pathway:

1) Neurons that were repetitively stimulated with 50 mM KCl in the presence of high concentrations of extracellular sucrose, a condition which inhibits clathrin-mediated endocytosis (Daukas & Zigmond, 1985; Heuser & Anderson, 1989) gave a significantly larger response to AMPA following KCl stimulation. The enhanced AMPA response consequent to KCl-induced depolarization could, however, be inhibited with botulinum toxin E, which proteolytically cleaves SNAP-25 (synaptosome-associated protein of 25 kD) or with tetanus toxin, which proteolytically cleaves the vesicular membrane protein, VAMP2 (also known as neuronal synaptobrevin or n-Syb).
2) Upon treatment with α-latrotoxin, which not only induces synaptic vesicle exocytosis but also prevents synaptic vesicle endocytosis in calcium-free cultured hippocampal neurons (Pennuto *et al.*, 2002), antibodies targeting the extracellular portion of the GRIA2 subunit gave a significant staining of the growth cone indicating that the massive fusion of synaptic vesicle was accompanied by the insertion of AMPA receptor subunits into the plasma membrane. The same is true for mature synapses. Synaptosomes purified from adult rat forebrain and stimulated with 0.1 nM α-latrotoxin in the absence of extracellular calcium showed an increase in cell surface GRIA2 (but not GRIN1) and synaptophysin, an integral synaptic vesicle protein. Furthermore, synaptophysin, VAMP2, GRIA2/3 and GRIA1 were co-enriched in a vesicular fraction immunoisolated using magnetic beads coated with antibodies directed against synaptotagmin. Immunolabeling of a highly purified synaptic vesicle fraction, prepared via permeation chromatography on controlled-pore glass, also revealed that GRIA1 and GRIA2 copurify with the synaptic-vesicle protein synapsin I.

In short, the data above indicates that AMPA receptor subunits reside in synaptic vesicle membranes and that these vesicles mediate the delivery of the AMPA receptor subunits to the plasma membrane, facilitated by stimuli (eg. α-latrotoxin) which promote synaptic

vesicle exocytosis. Interestingly, application of AMPA also affects the distribution of synaptic vesicles within hippocampal growth cones (Schenk *et al.*, 2003). Vesicle relocation to the tip of the growth cone filopodia could be detected upon AMPA application. Quantification of growth cone area immunoreactive to VAMP2 showed a two-fold increase in AMPA-treated cultures (Schenk *et al.*, 2003).

The experimental evidences above support the idea that AMPA receptors are presented to the cell surface by synaptic vesicles through an exocytic pathway involving the membrane fusion machinery proteins, NSF/α-SNAP and SNAREs, and other synaptic vesicle membrane proteins.

1.6 Aims of the present study

The primary aim of this thesis is to identify regulatory elements that are enriched within the promoters of the GRIA family of genes that control their expression. The identification of key transcriptional regulatory elements in the GRIA promoters has a greater global significance in that it can be applied in the design of novel gene-targeting constructs. For example, the identification of say, a neuron-specific glutamate receptor promoter element could possibly be used to deliver future experimental transgene and therapeutic agents to selected neurons in the brain.

To acheive the primary goal of this thesis, I first developed an algorithm for a software for the automated collection of human gene promoter sequences which I called "5'-end Information Extraction" or FIE. This is described in Chapter 2.

The data collected by FIE (version 2) or FIE 2.0 (Chong *et al.*, 2003), some 10,000-odd human gene promoters, were used by the program developed and described by Bajic *et al.* (2004) to find distinct transcriptional regulatory elements within the promoter sequences of the AMPA receptor subunits. In addition, a phylogenetic footprinting study of the GRIA1 subunit was also carried out to find TFBS sequences conserved through

evolution within the promoter region of the GRIA1s. The key promoter elements elucidated by these two approaches are described in Chapter 3.

In the course of this study, 47 genes were identified that shared the very combination of transcriptional regulatory elements found in the promoters of AMPA receptor subunits, by the method developed by Bajic *et al.* (2004). It was, thus, proposed that these 47 genes are co-regulated and/or co-expressed with AMPA receptors. To substantiate this claim, I studied in detail 7 of these genes and provide supporting evidence of how they may be involved in AMPA receptor expression and physiology in Chapter 4.

Due to the nature of the work, the methods and computational/software tools used will be described within each individual chapter, where applicable.

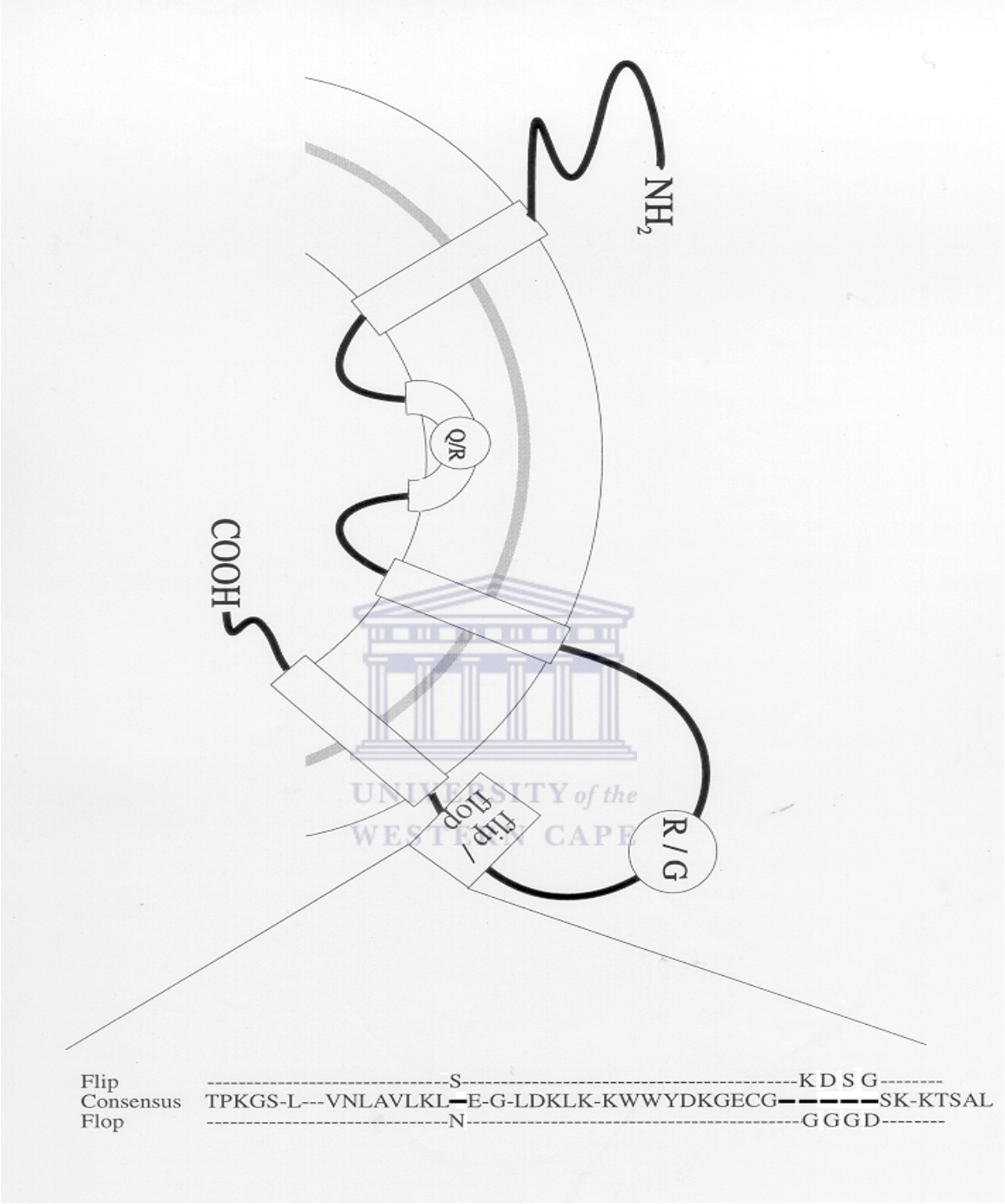


Figure 1.1:

Proposed secondary structure of glutamate receptor subunit depicting critical sites conferring functional diversity on AMPA receptors. The flip/flop splice cassette is shown with the consensus amino acid sequence and the residues which characterize the flip and flop variants.

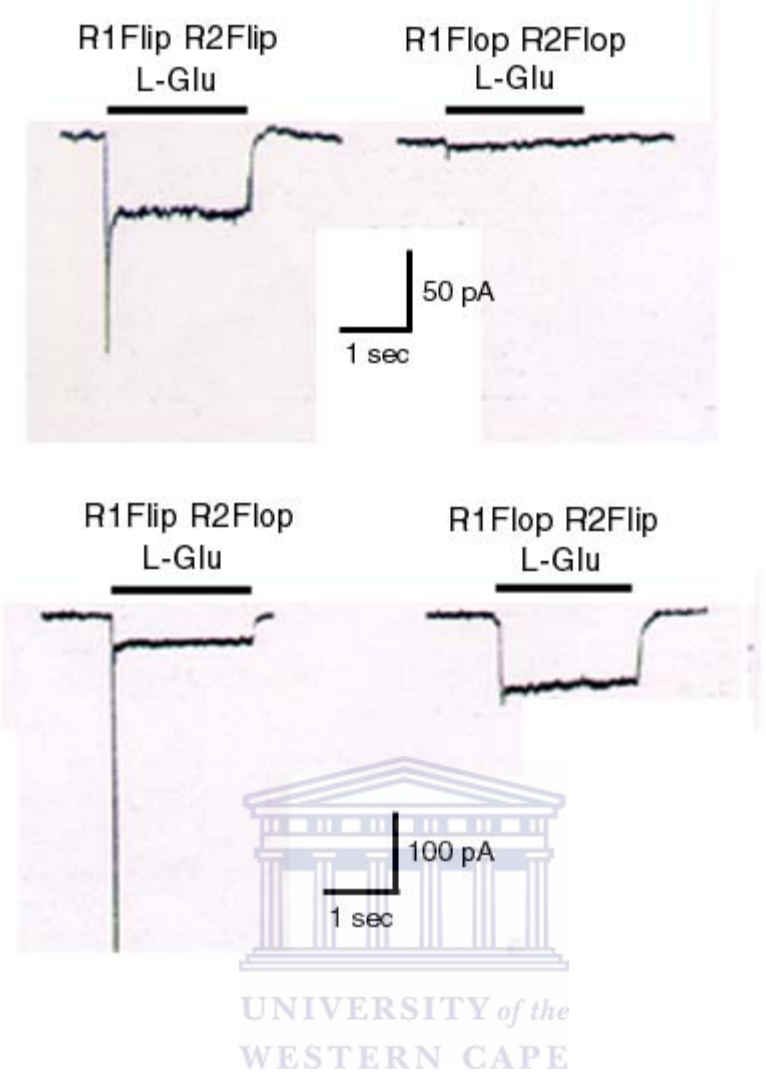


Figure 1.2:

Electrophysiological recording on recombinant glutamate receptors expressed transiently in human embryonic kidney cells 293. The receptors shown above are heteromers formed by flip and flop variants of the glutamate receptor subunits GRIA1 (denoted "R1" in the figure) and GRIA2 ("R2"). 300 μM of L-glutamate (L-Glu) is applied in each case. By simple deduction, it is easy to see that the flip variant is more effective and therefore, elicits a bigger response than its corresponding flop variants. Adapted from Sommer *et al.*, 1990.

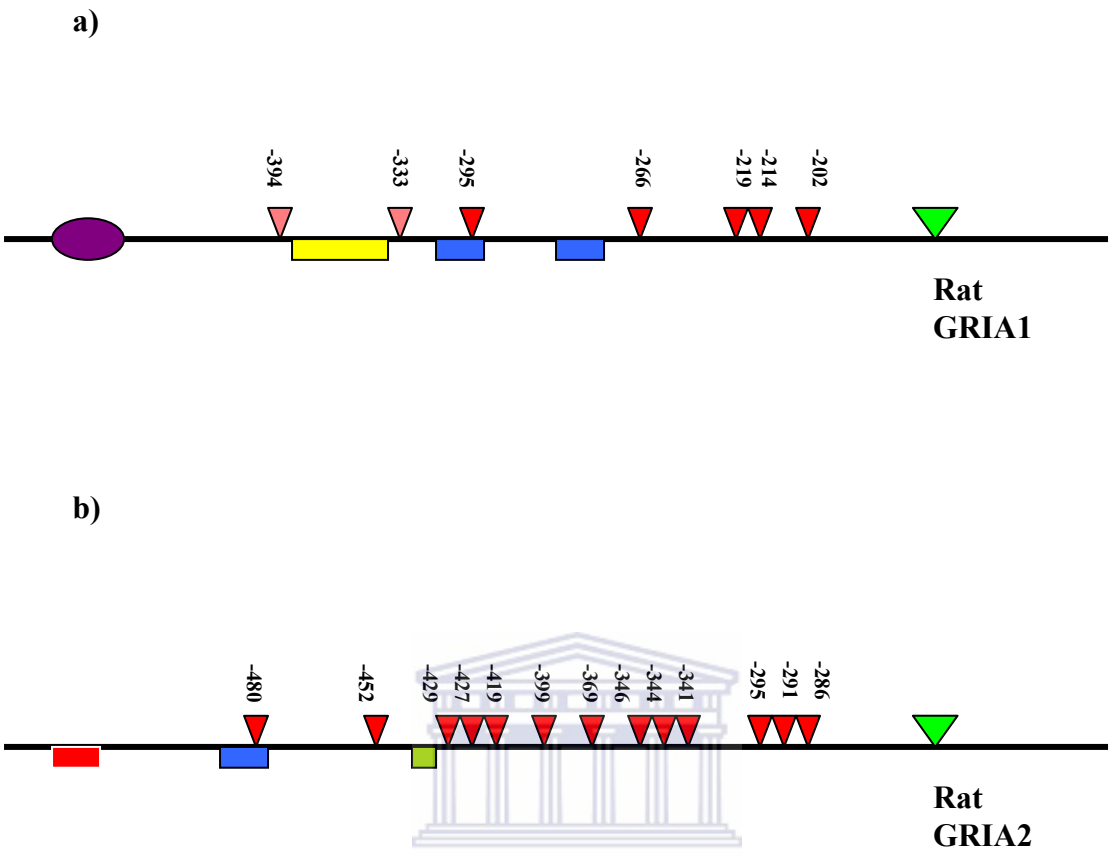


Figure 1.3

The above figure gives a diagrammatic view of the promoter regions of a) GRIA1 and b) GRIA2 rat genes and the transcriptional elements that have been elucidated experimentally. Both genes are shown to have multiple transcription start sites (inverted red triangles and pink triangles – TSS indicated by inverted pink triangles are weak transcription start sites). For the GRIA1 gene, two Sp1 binding sites (blue boxes) are found close to the TSSs at positions -296 and -275 while for GRIA2, one Sp1 binding site was found close to the 5'-most TSS at around -480. In GRIA1, a 64bp GA repeat is found upstream of the Sp1 binding site. In addition, a 57bp region containing an N-box (purple oval) is found at -743 and -686 in the GRIA1 gene. For the GRIA2 promoter, a neuron-restrictive silencer element (NRSE) (red box) is found about 200bp upstream of the 5'-most TSS and Sp1 binding site. Slightly downstream of the Sp1 binding site in the GRIA2 promoter is a nuclear respiratory factor-1 (NRF-1) binding site (green box). All positions are given with respect to the translation initiation site (inverted green triangle). Both GRIA1 and GRIA2 genes do not have any TATA or CCAAT boxes.

CHAPTER 2: DEVELOPMENT OF THE FIE SYSTEM FOR COLLECTION OF HUMAN GENE PROMOTER SEQUENCES

2.1 Aims

The best known collection of promoter sequences is the Eukaryotic Promoter Database (EPD) which is a carefully curated, nonredundant collection of experimentally-verified eukaryotic RNA Pol II promoters (Cavin Périer *et al.*, 1998; Praz *et al.*, 2002; Schmid *et al.*, 2004). At the start of this work in 2001/2002, Release 67 of the EPD contained only 1390 promoters, however, this collection included promoters of both multicellular plants and animals (Praz *et al.*, 2002). Traditionally, promoter sequences were obtained from the nucleotide databases of GenBank and EMBL (Benson *et al.*, 2005; Kanz *et al.*, 2005) and biological literature. Zhang & Zhang (2001) attempted to automate the process of extraction of promoter sequences from GenBank with the development of their PEG software. The extraction of promoter sequences from GenBank or EMBL records have serious limitations, most notably, the length of the promoter sequence is variable and is determined by the extent to which the respective labs have cloned the gene's promoter region.

In developing the *5*'-end *I*nformation *E*xtraction (FIE) system, the aim was to develop a system which could not only automate the collection of promoters but also, (1) to surpass the number of promoter sequences then available in the existing EPD and (2) to overcome the limitation of extraction methods that relied on GenBank records.

2.2 Introduction

Regulation of gene expression is mediated mainly through the promoter. Promoters are stretches of DNA sequences, generally located upstream of and overlapping the transcription start site (TSS) of genes. There is an abundance of mRNA/cDNA sequence information from public databases, such as GenBank, which are available to molecular

biology researchers and bioinformaticians for study. However, therein lies a key problem: while information is plentiful and readily available, the information may also be disparate and incomplete. For example, mRNA sequences for a particular gene may be of varying length because different labs who have attempted to clone the gene may have achieved this with varying degrees of success; therefore some of the mRNA sequences entered into GenBank may be 5'-incomplete. The solution is to try and find a way to filter out as much valuable information as possible from these public databases such that we may use all these sequence information to study the characteristics of the gene start region in order to gain a better insight into gene expression.

FIE (version 1) is the **first** software to rely on the principle of using the alignment of mRNA/cDNA sequences on the human genomic contigs to determine the TSS position of a gene on the contig and subsequently, using this information, to extract a user-defined sequence of the promoter region from the respective genomic contig (Chong *et al.*, 2002). Other similar programs that work on this principle include the PromoSer software (Halees *et al.*, 2003). The Evidence Viewer (EV) page of NCBI's LocusLink provides an alignment of a representative set of mRNAs/cDNAs on the human genomic sequence for each gene (Maglott *et al.*, 2000; Pruitt *et al.*, 2000, 2001). Both versions of FIE (Chong *et al.*, 2002, 2003), relies on this curated database for the alignment of mRNA/cDNA sequences on the human genomic contigs and the FIE system analyzes these alignments to determine the 5'-most end of gene on its respective chromosome. In this chapter, I shall give an account of the development of the FIE2 software (Chong *et al.*, 2003), the latest version of the FIE system that was used to extract some 10,000-odd human promoter sequences.

FIE2 is a specialized program for the extraction of genomic DNA sequences around the start (promoter region) and translation initiation site (TIS) of a gene. The start and TIS positions of the gene are determined from the alignment of a set of mRNAs representative of the gene of interest on the human genomic sequence, as given by LocusLink's EV page (Maglott *et al.*, 2000; Pruitt *et al.*, 2000, 2001). As a result of multiple alignments of mRNAs on the genomic sequence given on LocusLink's EV page,

multiple start positions are usually given for a gene. The 5'-most ***S***tart ***O***f ***E***xon ***1*** (SOE1) position identified by FIE2 thus represents the 5'-most position of the alignments of the representative mRNA transcript(s) on the genomic contig. At the time of publication of the FIE2 system, there were two other programs, Promoter Extraction from GenBank (PEG) (Zhang & Zhang, 2001) and EZ-Retrieve (Zhang *et al.*, 2002), which although are similar in their goal to FIE2, differ from FIE2 in functionality and methodology: primarily PEG draws its extraction from GenBank's mRNA records instead of the human working draft genomic sequences while EZ-Retrieve uses the Abstract-Syntax-Notation-One (ASN.1) files to get an approximate position of the gene's start which is not always supported by gene transcripts. Furthermore, both PEG and EZ-Retrieve cannot extract sequences around the TIS and PEG is also not accessible from the Web.

The importance of the sequences extracted by FIE2 also lies in its usefulness for follow-up experiments in the lab in current research efforts to understand the transcriptional machinery. In addition, the sequences can also be used to compile datasets for training and testing gene finding/prediction systems, such as Dragon Promoter Finder (Bajic *et al.*, 2002a, b) and Dragon Gene Start Finder (Bajic & Seah, 2003).

2.3 Program Description

FIE2 can be accessed at the URL address: http://research.i2r.a-star.edu.sg/FIE2.0. The web interface for FIE2 is fairly intuitive. Input to FIE2 can either be a gene or protein name or LocusID (for additional query options, please refer to LocusLink's help page: http://www.ncbi.nlm.nih.gov/LocusLink). Users must also input the length of sequence upstream and downstream of the start of the gene (which is abbreviated, SOE1 ['start of exon 1'], by FIE2) that they wish to extract. The user is encouraged to be as specific as possible when submitting a query to FIE2; for example, where possible users should use the option to search by LocusID. If a general query (e.g. actin) is submitted to FIE2, this is sent to LocusLink which then returns a list of links to genes which match the query. The user must choose the appropriate link for the gene of his interest. The user's request is again submitted to LocusLink which then returns an information page which is

processed by FIE2. Among the information that FIE2 gathers from LocusLink's information page is the availability of an EV page. For human genes, LocusLink attempts to align a set of published sequences representative of a gene on its respective chromosomal sequence (genomic contig) in its EV page (Figure 2.1a). The number and specific instances of accession numbers (gene records) used in the alignment depend on whether the gene has a provisional or reviewed reference sequence (RefSeq, http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html) record, or no RefSeq record at all. If no EV page is available, FIE2 returns a "No Evidence Viewer found" page. If the initial query submitted to FIE2 is very general (e.g. actin), LocusLink would return a fairly long list of gene links. Unfortunately, there is no way for FIE2 to determine if the gene links contain an EV page without making several hits to the LocusLink server to gather the LocusLink information page for each and every one of those genes on the list (this is an undesirable practice since it places an unnecessary burden on the LocusLink server). One may convert accession numbers to LocusID values using the daily updated file that is available from ftp://ncbi.nlm.nih.gov/refseq/LocusLink/loc2acc.

It is possible that the LocusLink EV page presents more than one gene along a genomic region of the contig. FIE2 attempts to recognize all relevant (valid) accessions (mRNA sequences) by gene name or symbol or other aliases among the accessions presented in the sequence alignment on the EV page. The valid accessions are abbreviated 'GD' or 'AA' by FIE2 depending on whether the accession was identified as valid based on a match of its gene name/description/symbol or an alias/alternative symbol, respectively (Figure 2.1c). Based on the sequence alignment given on the EV page, if an accession has a high sequence identity with the 5'-end exon of an already identified valid accession (but was not previously recognized as a valid accession by its gene name or symbol), then FIE2 labels this as an 'AVA' (associated valid accession) (Figure 2.1a & 2.1c). The gene description of the AVA is given alongside its accession on the result page bearing the SOE1 position information (*no descriptions are given for valid accessions, GD and AA*) (Figure 2.1b). The SOE1 positions, based on the alignment of all valid accessions and AVAs on the genomic contig, are presented to the user for his analysis. The FASTA sequence (with the user-specified length) around all the SOE1 positions identified by

FIE2 can be retrieved through their respective ‘View FASTA Sequence’ hyperlink (Figure 2.1b). The 5’-most SOE1 position identified by FIE2 thus represents the 5’-most position of the alignments of the representative mRNA transcript(s) on the genomic contig.

Along with the DNA sequence alignment on the genomic contig, the EV page also presents an alignment of the coding sequence of each accession alongside its DNA sequence. For each valid accession, FIE2 locates the position of the TIS on the genomic contig by identifying the position of the start of its coding sequence (Figure 2.1a). If the TIS is found to be in ‘Exon 1’ of the genomic region presented on the EV page then the SOE1 position for each valid accession or AVA is identified individually based on the position of 5’-most end of the aligned mRNA sequence on the genomic contig. For example, in the case of RABL4, we can see that the start of the coding sequence is in ‘Exon 1’(Figure 2.1a) and the SOE1 positions based on the 5’-most position of NM_006860 and BC000566 on the contig are identified by FIE2 and displayed accordingly, as shown in Figure 2.1b. If, however, the TIS position does not meet the above criteria, then in addition to identifying the SOE1 position for each valid accession on the genomic contig, a process is also initiated to determine if ‘Exon 1’ presented on the EV page is indeed the first exon of the queried gene. This is because in certain cases, where there is more than one gene in the genomic region presented on the EV page, the first exon presented in the sequence alignment might not represent the first exon of the queried gene. In such instances, the true first exon is identified by locating the exon containing the 5’-most position of all the aligned mRNAs of the queried gene. FIE2 then renumbers all exons on the EV page accordingly so as to reflect the true exonic-intronic partition of the gene sequence on the genomic contig. This step is crucial in determining the exon(s) that contains the SOE1 and TIS.

For FIE2, multiple SOE1 positions may be presented (as explained above) and likewise, multiple TIS positions may also be given. The given coding sequence for some of the valid accessions or AVAs may sometimes be predicted and therefore, the position of the TIS differs from those of other valid accessions used to align against the genomic contig.

The coding sequence and thus, the TIS is sometimes predicted and not experimentally verified by the lab which cloned the cDNA sequence; however, such information/annotation is not provided by LocusLink. Therefore, FIE2 presents all TIS positions (predicted or otherwise) for the sake of completeness. In some cases, the presence of multiple TISs may also be due to the different initiation sites for different transcript variants.

Several scenarios may present themselves that might leave FIE2 to tag the given SOE1 position as "indeterminate" (that is to say, FIE2 cannot determine the SOE1 position). Four categories of "indeterminate" exist in FIE2 and these are explained as such:

- **Indeterminate 1**: The coding sequence for all aligned mRNAs of the gene of interest indicates that the start codon lies upstream of their 5'-most aligned position on the genomic contig, and therefore, also making the position of the start of exon 1 (SOE1) indeterminate.
- **Indeterminate 2**: More than one translation initiation site (TIS) was identified for this particular gene, however, the coding sequence for one (or more) of the aligned mRNAs indicates that the start codon is indeterminate. From experience, the developers of FIE have found that these aligned mRNA sequences with indeterminate TIS positions are sequences belonging to cDNAs clones from the German Cancer Research Center (DKFZ), Kazusa DNA Research Institute (KIAA or FLJ: see http://www.kazusa.or.jp/NEDO/ for an explanation of the differences between KIAA and FLJ sequences), or IMAGE Consortium. The sequences of these clones have had their coding sequence predicted and/or were annotated to have a partial coding sequence, and thus, the position of their start codon was not experimentally identified. Therefore, in all likelihood, the SOE1 position identified by FIE2 may represent a true SOE1 position.
- **Indeterminate 3**: None of the mRNAs used in the alignment against the genomic contig on the EV page were identified as an exact match for the gene of interest based on the gene symbol, description, alternate symbol or alias of the gene. The SOE1 position given by FIE is probably based on an alignment of an mRNA sequence that is highly similar to the gene of interest or possibly from a cDNAs clones from the

German Cancer Research Center (DKFZ), Kazusa DNA Research Institute (KIAA or FLJ: see http://www.kazusa.or.jp/NEDO/ for an explanation of the differences between KIAA and FLJ sequences), or IMAGE Consortium.. Therefore, this SOE1 position is viewed as indeterminate.

- **Indeterminate 4**: Although there is at least one aligned mRNA sequence that is an exact match for the gene of interest, its coding sequence is not provided on the EV page. Therefore, FIE is unable to determine if this mRNA sequence is 5' complete. The SOE1 position is therefore considered to be indeterminate.

FIE2 still retrieves the sequence upstream and downstream of the 5'-most position of the mRNA alignment on the genomic contig in these cases, but with the caveat that the SOE1 position is 'indeterminate'. It is up to the user's discretion to determine whether or not to use the sequence provided.

FIE2 also determines the strand orientation of the gene on the genomic contig. If the locus is found on the complementary strand of the contig, FIE2 retrieves the user-specified sequence region and presents the FASTA sequence in its reverse complement. A similar process is carried out to retrieve the FASTA sequence around the TIS where available. The header for the FASTA sequence can be interpreted as follows: for example, for LocusID 10043, if the user chooses to extract 10 bp upstream and downstream of the SOE1 position, then the header would read ">gi|16168698:14992135-14992154 Homo sapiens chromosome 22 reference genomic contig|SOE1|14992145". This mean that the SOE1 for this gene (in this case, based on the alignment of AK026576) in found in the genomic contig (GI [NCBI's sequence identifier]: 16168698) for chromosome 22 and the SOE1 position is 14992145 on this contig. The 20 nt. long sequence requested by the user therefore stretches from position 14992135-14992154 on the contig.

The following additional information on the gene of interest is also provided:

1. the descriptive name of the gene
2. alternate symbols / aliases for the gene

3. the chromosome on which this genetic locus is found
4. the gene's cytogenetic position on the above chromosome
5. the accession number for the genomic contig on which this locus is found
6. the GI (GenBank's unquie identifier) number for the contig

An example of the output returned by FIE2 can be seen in Figure 2.1b.

Flowcharts depicting FIE2.0's algorithm are given in Appendix 1.

2.4 Results

2.4.1 Testing FIE2

An updated annotation for human chromosome 22 was released when development of FIE2 reached its final phase (Collins *et al.*, 2003). Therefore, it was decided that the new anotations for chromosome 22 would be used to benchmark FIE2. There were altogether 393 mRNA sequences of protein coding genes which were considered complete and mapped to the genomic sequence for human chromosome 22 (http://www.sanger.ac.uk/cgi-bin/hgp/chr22/display?Chr22.3.1b.coding_genes.gff). Although not all known genes from the current Sanger chromosome 22 annotation were yet included in LocusLink, 230 of these genes could still be found in LocusLink. For these 230 genes, FIE2 could determine the SOE1 position for 208 and the SOE1 position for the remaining 22 genes were either tagged as "indeterminate" or given a tag that represented the fact that there was no EV page or an incomplete EV page for that particular LocusID.

Table 2.1 shows the distribution of SOE1 positions when compared against Sanger's annotated gene start positions. Of the 208 genes whose SOE1 position could be determined by FIE2, 40 matched EXACTLY the gene start positions annotated by Sanger, while 54 were, in fact, found to be UPSTREAM of the positions given by Sanger. The SOE1 position extracted by FIE2 for the remaining 112 genes were all found

to be downstream of Sanger's annotated positions. Even so, the SOE1 positions of 74 of the remaining 112 genes were within 100 bp of the annotated position given by Sanger. While 80.8% (168 of 208 genes) of the extracted SOE1 positions were either accurate or within 100 bp of current annotations given by Sanger, 12% (25 genes) are between 100 and 1,000 bp from the annotated positions and ONLY 6.3% (13 genes) are beyond 1000 bp downstream of the annotated position. Two anomalies were also found: ARFGAP1 and GSTT2. ARFGAP1 was mapped to Chromosome 20 by LocusLink while in the case of GSTT2, the SOE1 position retrieved from LocusLink and the gene start position annotated by Sanger differed by 18,931nt. in length. In addition, the gene orientation of GSTT2 on the genomic sequence given by LocusLink was also different from Sanger's current annotation.

Figure 2.2 gives a histogram of the distribution of SOE1 positions that were given downstream of Sanger's annotated gene start positions. The full result of the extraction for all 230 genes is given in Appendix 2. A detailed table giving a comparison of SOE1 positions extracted by FIE2 for each individual gene against annotated gene start positions from Sanger is also given in Appendix 3. Although the positions extracted by FIE2 are given relative to the genomic contig, these contig positions were converted to chromosomal position for these 208 genes for easy reading. The calculations are based on information given at NCBI of the contig-to-chromosome positions.

2.5 Discussion

2.5.1 FIE2: the program and its capabilities

FIE2 has proven to be more effective than FIE (version 1) (Chong *et al.*, 2002) in extracting accurate information on the SOE1 and TIS of a gene. This improved performance is due mainly to its ability to filter out irrelevant gene sequence alignments

on LocusLink's EV page when more than one gene is aligned in a genomic region. An added feature of FIE2 is its ability to provide users with multiple SOE1 positions based on alignment of various mRNA transcript sequences on the chromosomal sequence. RefSeq along with other supporting sequences are used to verify the genetic locus on the contig in LocusLink. However, the authors of DBTSS (database of transcriptional start sites: http://dbtss.hgc.jp/samp_home.html), using the oligo-capping method for creating cDNA libraries, found that about a third of the RefSeqs are not 5'-end complete (Suzuki *et al.*, 1997, 2002; Yamashita *et al.*, 2001). The current version of LocusLink does, in fact, include the use of full-length cDNAs from the NEDO human cDNA sequencing project (denoted by the prefix "FLJ") (generated by the "oligo-capping" method) (Maruyama & Sugano, 1994; Sugano: http://www.nedo.go.jp/bio-e/index_syokai.html) and large cDNAs (> 4 kb) of the Kazusa human cDNA sequencing project (denoted by the prefix "KIAA") (generated by conventional methods) (Ohara *et al.*, 1997; Nagase *et al.*, 2001). In addition, cDNA sequences from the German Cancer Research Center (DKFZ) (Wellenreuther *et al.*, 2001; Wiemann *et al.*, 2001) and IMAGE Consortium (Lennon *et al.*, 1996) are also used by LocusLink. A good number of the FLJ, KIAA, DKFZ and IMAGE cDNAs are uncharacterized and FIE2 needs to make an "educated guess" as to whether these sequences represent the gene in question. If these cDNA clones bear some identity to the 5'-end of a known sequence of the gene, the sequence is considered to represent the gene in question. This cDNA sequence is then labeled as an 'AVA'. The 5'-most aligned position of this AVA sequence on the contig is then given as a suggested SOE1 of the gene (Figure 2.1a & 2.1b).

2.5.2 An analysis of FIE2's accuracy

New gene annotations were released for human chromosome 22 by the Wellcome Trust Sanger Institute (Collins *et al.*, 2003). Under these new annotations, there are 393 mRNA sequences of protein coding genes which are considered complete and mapped to the genomic sequence for chromosome 22 (http://www.sanger.ac.uk/cgi-bin/hgp/chr22/display?Chr22.3.1b.coding_genes.gff). To provide a benchmark for FIE2, a retrieval of SOE1 information using FIE2 was performed and FIE2's results were

compared with the new annotations from Sanger. Using the names in the “Locus” field of the Chr22.3.1b.coding_genes.gff file, a search was made for their corresponding entries in LocusLink. Of the 393 “complete” genes annotated by Sanger, only 230 matched entries in LocusLink. The remaining 163 “complete” genes annotated by Sanger were named with an accession number which was not recognized by LocusLink, for example, “Em:AC005500.C22.3” (with the description "Matches EST sequences"). Therefore, only these 230 “complete” genes were used to gauge FIE2’s effectiveness against the current available annotations from Sanger. FIE2 was able to determine the SOE1 positions for 208 genes based on LocusLink’s EV page. The SOE1 positions for the remaining 22 genes were either tagged as “indeterminate” by FIE2 or could not be retrieved because there was no EV page or an incomplete EV page. Comparing the SOE1 positions of the 208 genes with the new annotations from Sanger, it was found that:

1. For 40 genes, the 5’-most SOE1 position identified by FIE2 were identical to current annotations from Sanger
2. The 5’-most SOE1 position for 54 genes was extended upstream of the annotated Sanger gene start position
3. The 5’-most SOE1 position for 112 genes was found to be downstream of the annotated Sanger gene start position
4. The information retrieved by FIE2 for 2 genes (ARFGAP1 and GSTT2) did not agree with Sanger’s annotations

The search results for all 230 “complete” genes can be viewed in Appendix 2.

In the case of the 2 genes, ARFGAP1 and GSTT2, it was found that ARFGAP1 had, in fact, been mapped to Chromosome 20 by LocusLink. As for GSTT2, the SOE1 position retrieved by FIE2 and the gene start position annotated by Sanger differed by 18,931nt. in length *and* the gene orientation on the genomic sequence was also in contention.

For those 112 genes whose 5’-most SOE1 positions were downstream of annotated Sanger gene start positions, it was found that the difference between FIE2’s and Sanger’s position did not exceed 100nt. in length for 74 genes. For 20 genes, the 5’-most SOE1

position were found to be downstream by between 100-500nt. of Sanger's annotated gene start position. Therefore, 90.4% (188 of 208 genes) of the extracted SOE1 positions were either accurate or within 500 bp of current annotations given by Sanger.

For the genes which had their 5'-most position extended upstream of the Sanger's annotated gene start position, these extensions were based *either*:

on RefSeqs or other representative mRNA sequences of the gene of concern:-
For example, the SOE1 position for MPST (LID: 4357) could be extended 4,395nt. upstream of Sanger's annotated position based on the alignment of its RefSeq sequence, NM_021126 on the genomic sequence;

or

on 'AVA's which bore a high degree of identity to the gene of concern in the 5'-end of the sequence:-
For example, in the case of ARHGAP8 (LID: 23779), the 5'-most position given by FIE2 is based on an alignment of a NEDO sequence, AK091884 (an 'AVA'). This AVA bears a high degree of identity to 2 mRNA sequences for the gene, AF195968 and AK000192 (defined as 'FLJ20185', an alias for 'ARHGAP8'), over their entire length. The resulting use of the AVA, NEDO sequence AK091884, extended the 5'-most SOE1 position by 25,197nt. upstream of Sanger's start position. The description of the AVA is given alongside its accession on FIE2's information page, in this case: "AK091884 (Homo sapiens cDNA FLJ34565 fis, clone KIDNE2006210, moderately similar to Rho GTPase activating protein 8)". It is therefore up to the user to evaluate as whether these AVAs are extended sequences of the gene of interest, perhaps a yet-to-be-characterized transcript variant, or an entirely different gene.

However, in each case, the user can be assured that the SOE1 position is determined as a result of the alignment of an *experimentally derived* mRNA sequence against the genomic sequence.

I did, in fact, find that the 5'-most SOE1 position for BCR (LID: 613) was wrongly extended upstream of the annotated Sanger's position. The 5'-most SOE1 position was wrongly identified by FIE2 because the EV page contained an alignment using M64437. M64437's GenBank record defines it simply as "Human BCR mRNA, 5' end". The M64437 sequence is actually a sequence containing the BCR promoter (Shah *et al.*, 1991). Therefore, this BCR sequence begins beyond and upstream of the BCR TSS. The 5'-most SOE1 position mistakenly identified by FIE2 extended the 5'-end by 155nt upstream. However, it should be noted that the correct SOE1 position was also identified by FIE2, based on the alignment of the mRNA sequences, NM_021574 and Y00661, on the contig and this is IDENTICAL to the annotated Sanger gene start position.

The main reasons for FIE2's inability to extract information of the SOE1 or TIS position is usually due either to the lack of an EV page in LocusLink (that is to say, no sequence alignment on the respective genomic contig was carried out for the gene in question), or to the 5'-end of the gene sequence falling within a gapped region of the chromosome (a region where the genomic sequence has not been elucidated).

FIE2's name recognition ability has been greatly enhanced (over FIE version 1) as certain adjectives / terminologies are no longer recognized by FIE2 as being part of the gene name or symbol. However, its name recognition module can be further improved to recognize subtle nuances without comprising on its speed. For example, given the gene name LIMK2, FIE2 is currently unable to tell that the names, LIMK-2 / LIMK 2, also represent the sequence.

Furthermore, FIE2 is also able to class SOE1 positions deemed to be indeterminate into four separate categories (a detail explanation of these 4 categories is given above in section 2.3). Such detailed classification was previously not provided by FIE version 1.

2.5.3 Comparison of FIE2 against other similar programs

A program for the extraction of eukaryotic promoter sequences from GenBank (abbreviated to PEG), was developed by Zhang & Zhang (2001). The similarities and differences between the FIE2 and the PEG programs are as follows:

1. Multiple SOE1 positions are presented by FIE2 based on the determination of the position of the 5'-most end of various annotated mRNA sequences which are deemed to be representative of a gene. The set of representative mRNA sequences are preselected by NCBI's LocusLink and may sometimes include full-length, uncharacterized cDNA sequences from the NEDO human cDNA sequencing project (FLJ), Kazusa DNA Research Institute (KIAA), German Cancer Research Center (DKFZ), and IMAGE Consortium, which are highly similar to the gene in question. These FLJ, KIAA, DKFZ and IMAGE sequences are usually denoted by an accession number assigned by the individual research institute. PEG searches the 5'-most mRNAs of the gene of concern by iteratively extending mRNA sequences at the 5'-end and it is possible that such cDNA sequences (FLJ, KIAA, DKFZ, IMAGE) could be omitted by PEG.

2. Sequences extracted by PEG can only go as far upstream as is annotated in GenBank's record, and thus cannot be directly extended further upstream. FIE2 does not have this limitation since the sequence extraction in our program is based on currently available human genomic sequences.

3. FIE2 is able to identify the TIS position of a gene and extract the sequence around it, while PEG does not have this functionality. In some cases, recognition of multiple TIS positions is possible where transcript variants for a particular gene are identified (for example, the gene ADSL [LocusID: 158] on chromosome 22 has 2 possible open reading frames [ORFs]: GI:28904 & GI:28905).

4. Currently, FIE2 only supports the extraction of human sequences, but PEG can extract sequences from a broader spectrum of organisms (eukaryotes).

5. Both PEG and FIE2 attempt to extract the promoter region based on currently available mRNA sequences – in the case of PEG, it does so from GenBank's records, while for FIE2, it does so based on curated RefSeq and other supporting mRNA sequences which LocusLink has identified and aligned against the genomic contig.

It has to be highlighted that, for both FIE2 and PEG programs, there is a possibility that the 5'-end of the mRNA sequence may be incomplete but that does not negate the importance of the information extracted by these two programs. Both PEG and FIE2 try to make *the best use of currently available information*. Although the methodology and functionality of the two programs might differ, the aims of both programs are similar: to try and extract a length of sequence around what might be the promoter region based on currently available information so as to help facilitate in follow-up experiments in the lab and *in silico* in the studies of gene expression regulation.

The TSS is usually a good reference marker of the promoter region and it is true that only a handful of TSSs have been experimentally verified, as annotated by the EPD (Cavin Périer *et al.*, 1998; Praz *et al.*, 2002; Schmid *et al.*, 2004). However, both FIE2 and PEG are *not trying to pinpoint the TSS*, but are, instead, trying to extract a length of sequence that contains all, or part, of the promoter region (in FIE2, this depends on the length specified by the user). The promoter region can cover a region upstream of and overlapping the TSS and perhaps, extending downstream, nearing the TIS.

Theoretically, the SOE1 ('start of exon 1') is the TSS. However, in FIE, the annotation "SOE1" is used loosely because the position, as given on LocusLink, may not sometimes be the true TSS but rather the 5'-most aligned position of an mRNA sequence on the genomic contig. For example, the mRNA sequences for the gene of concern may be 5'-incomplete or the alignment of mRNA sequences on the genomic sequence may not always provide a match with high identity in the 5'-end. Thus, the 5'-most position of the alignment on the genomic sequence may not represent the true starting point of exon 1.

A second program, EZ-Retrieve, aims to retrieve the promoter region of genes using LocusLink's Abstract-Syntax-Notation-One (ASN.1) annotation file to obtain the gene's coordinates on the contig (Zhang *et al.*, 2002). However, this gives only an approximation of the gene start position since the start coordinate given in the ASN.1 files refers to the locus and not the gene because LocusLink is, after all, locus-oriented. Two key differences between FIE2 and EZ-Retrieve can be summerized as follows:

1. FIE2 gives the users multiple SOE1 positions based on the alignment of a set mRNA sequences, in some cases, transcript variants representative of a gene whereas EZ-Retrieve identifies the approximate start position of a gene based on the locus coordinates and presents the user with a single approximate gene start position.

2. FIE2 is able to identify the TIS position for a gene whereas this feature is not available in EZ-Retrieve.

I also carried out a similar extraction with FIE (version 1) to make a fair comparative analysis of the original program with the current FIE2 program. I found that FIE (version 1) could only locate the SOE1 positions for 201 genes. However, on closer look, I found that the SOE1 positions for 19 out of the 201 genes differed from those extracted by FIE2. For example, the SOE1 position for ECGF1 (LocusID: 1890) wrongly identified at position 105,989 on the contig NT_011526.4 by FIE version 1 when, in fact, the 5'-most position should have been at position 105,455 based on an alignment of NM_001953 on the contig. This is due to the fact that FIE (version 1) takes the 5'-most position of the contig presented on the EV page when the EV page states that there is "1 gene found in this genomic region". It should be remembered that FIE version 1 will only process the EV page if there is only one gene in the genomic region presented (Chong *et al.*, 2002). In cases where the EV page states that there is more than 1 gene in the presented genomic region, then FIE (version 1) does not attempt to extract any information of the SOE1 or TIS position (Chong *et al.*, 2002). However, in the case of ECGF1, in spite of the fact that the EV pages states that there is only "1 gene found in this genomic region", we see that an mRNA sequence, S72487 (GenBank description: orf1 5' to PD-ECGF/TP...orf2 5'

to PD-ECGF/TP [human, epidermoid carcinoma cell line A431, mRNA, 3 genes, 1718 nt].), was also aligned to the contig. FIE2 correctly identified NM_001953 as a valid accession and therefore gave the correct 5'-most SOE1 position while also correctly identifying that S72487 was an invalid accession based on its ability to recognize the gene name or symbol in the mRNA's GenBank description. In summary, FIE version 1 was only able to identify correctly the SOE1 position for 182 genes as compared to FIE2's 208 genes (out of the possible 230 genes found on LocusLink). The full results of the extraction carried out using FIE version 1 can be viewed online at http://research.i2r.a-star.edu.sg/FIE/test-dataset.html.

There are three major human genome online resources, namely, NCBI's Human Genome information resource (http://www.ncbi.nlm.nih.gov/genome/guide/human/) (Wheeler *et al.*, 2002), Ensembl's Human Genome Browser (http://www.ensembl.org/Homo_sapiens/) (Hubbard *et al.*, 2002) and the Human Genome Browser at UCSC (http://genome.ucsc.edu/) (nicknamed "GoldenPath") (Kent *et al.*, 2002). The interactive tools offered by these organizations allow researchers to view the genome at a 'macroscopic' level – that is, at the level of an exon-intron, a gene or a chromosomal band (as opposed to a base-by-base level). FIE2 complements these browsers by giving researchers a tool for easy extraction of the base sequence of specific genomic regions around a gene's 5'-end. With Ensembl's Human Genome Browser, a user may search for a gene and get back such information as its genomic location, similarity matches (that is, related records pertaining to the gene in HUGO, SWISSPROT, etc.), transcript structure and protein structure. Opening another window on their computer, a user could, in theory, use Ensembl's EnsMart or ContigView to retrieve a customized length of DNA sequence and thus, seemingly perform the same functions as FIE2. However, it would be prudent to take note of the fact that although the genomic location for a particular gene is supported by comparisons to protein, cDNA and EST data, the given start coordinate of the gene is sometimes either a GeneWise or Genscan prediction. FIE2 'reads and interprets' the sequence alignments of representative mRNAs on the contig and then extracts and presents all information, based on its analysis, in a concise form. In effect, FIE2 provides an extension of LocusLink by

streamlining the extraction of genomic sequence around a gene's 5'-end. In some cases, FIE2 refines and reorganizes the LocusLink data to supply the user with more reliable information. This one-shot analysis and processing by FIE2 thus helps the user save valuable time and effort.

Coleman and colleagues (2002) estimated that the alignment of reference mRNAs to genomic sequence allows promoters to be identified for at least 75% of genes. This, therefore, lends support to the concept on which FIE2 in based on. The results for FIE2 are very promising and show definitively that the new algorithm for FIE2 is a vast improvement over that of the older version of FIE.

Table 2.1:

Summary of extracted SOE1 positions from FIE2 relative to the annotated gene start positions for 168 genes of Chromosome 22 given by the Sanger Institute.

Summary of results

	(a) No. of genes	**(b)** % of compared sequences	**(c)** % of total query
Extracted SOE1 position by FIE2 is the same as annotated position given by Sanger	40	19.2	17.4
Extracted SOE1 position by FIE2 is upstream of annotated position given by Sanger	54	26	23.5
FIE2's SOE1 position is downstream of annotated position given by Sanger			
<= 100bp	74	35.6	32.2
>100bp and <=500bp	20	9.6	8.7
>500bp and <=1,000bp	5	2.4	2.2
>1,000bp	13	6.3	5.7
Anomalous annotations found between Sanger and LocusLink	2	0.9	0.7
Total	208		

Key: Column (a) gives the number of genes whose extracted 5'-most SOE1 positions were found to be either identical or upstream or downstream to the annotated Sanger gene start positions. Column (b) gives the number of genes as a percentage of all 208 genes that were compared with Sanger's annotations. Column (c) gives the number of genes as a percentage of all 230 genes that were submitted to FIE2 for extraction.

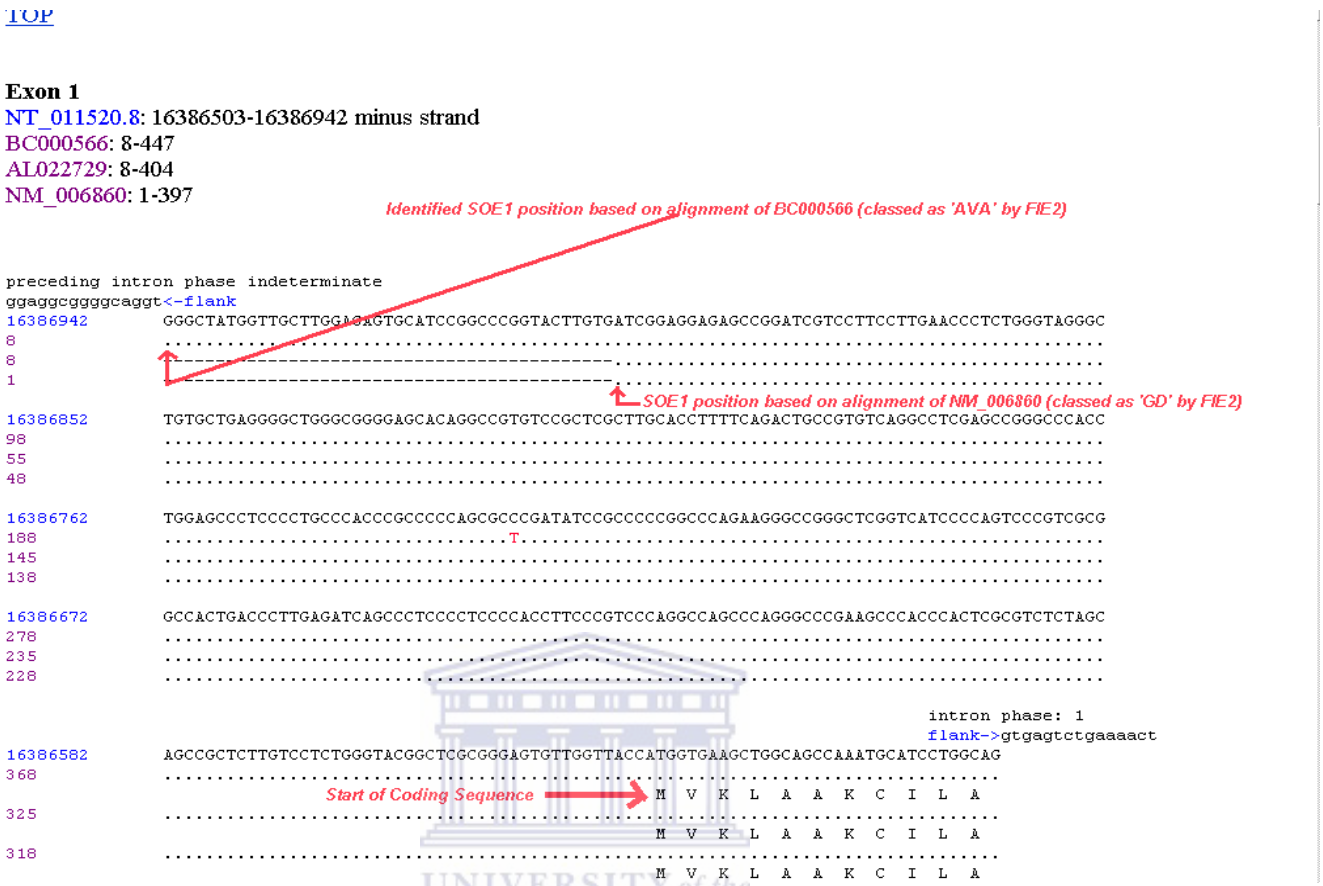
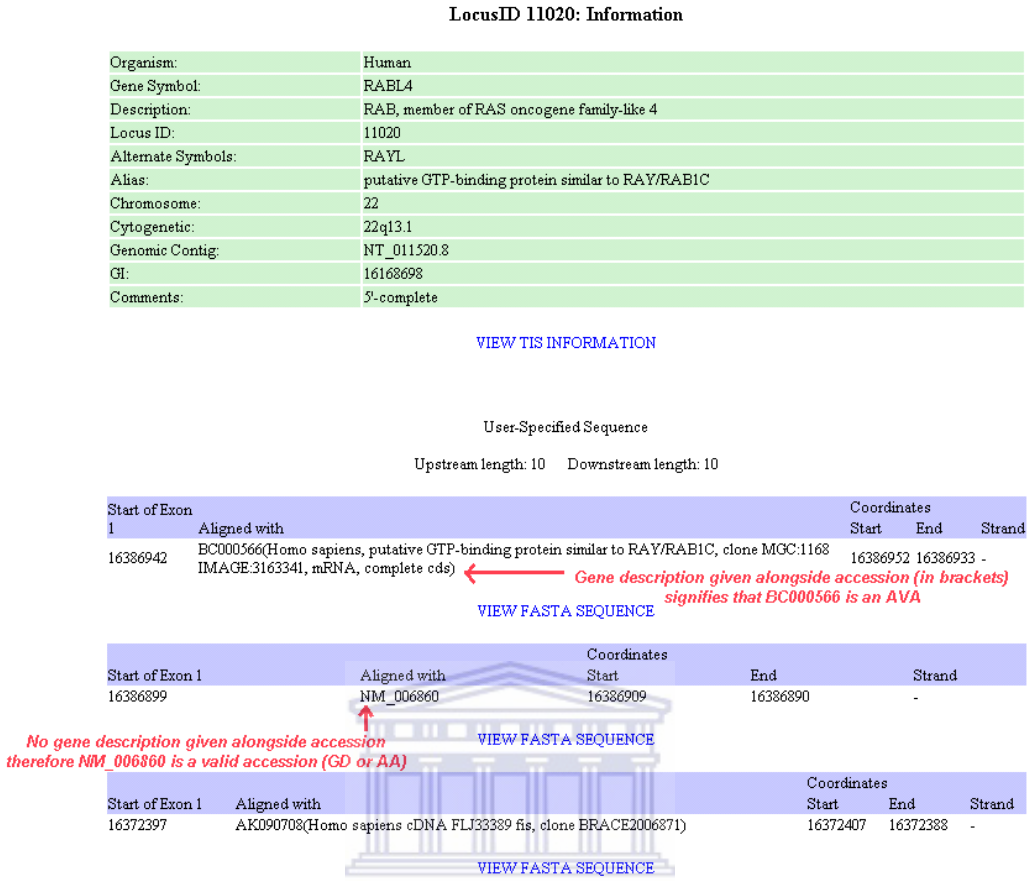


Figure 2.1a:

An example of LocusLink's EV page. In this EV page, the alignment of representative mRNA sequences for the gene RABL4 (LocusID: 11020) is shown. NM_006860 is recognized as a valid accession for the gene RABL4 through its exact match of it gene name/description and therefore classed as 'GD' (a valid accession) by FIE2. BC000566 (an IMAGE Consortium cDNA clone) is recognized as a possible mRNA sequence representative of the gene RABL4 because of its high degree of identity to the known RefSeq sequence for RABL4 (NM_006860). BC000566 is therefore classed as an associated valid accession or 'AVA' by FIE2. As can be seen here, the start of the coding sequence is found in 'Exon 1' of the genomic region present on the EV page. All labels in red were added by the authors for clarity of understanding.



LocusID 11020: Information

Organism:	Human
Gene Symbol:	RABL4
Description:	RAB, member of RAS oncogene family-like 4
Locus ID:	11020
Alternate Symbols:	RAYL
Alias:	putative GTP-binding protein similar to RAY/RAB1C
Chromosome:	22
Cytogenetic:	22q13.1
Genomic Contig:	NT_011520.8
GI:	16168698
Comments:	5'-complete

VIEW TIS INFORMATION

User-Specified Sequence

Upstream length: 10 Downstream length: 10

Start of Exon 1	Aligned with	Coordinates Start	End	Strand
16386942	BC000566(Homo sapiens, putative GTP-binding protein similar to RAY/RAB1C, clone MGC:1168 IMAGE:3163341, mRNA, complete cds)	16386952	16386933	-

Gene description given alongside accession (in brackets) signifies that BC000566 is an AVA

VIEW FASTA SEQUENCE

Start of Exon 1	Aligned with	Coordinates Start	End	Strand
16386899	NM_006860	16386909	16386890	-

No gene description given alongside accession therefore NM_006860 is a valid accession (GD or AA)

VIEW FASTA SEQUENCE

Start of Exon 1	Aligned with	Coordinates Start	End	Strand
16372397	AK090708(Homo sapiens cDNA FLJ33389 fis, clone BRACE2006871)	16372407	16372388	-

VIEW FASTA SEQUENCE

Figure 2.1b:

An example of FIE2's SOE1 Information page. Here, we see that 3 SOE1 positions are given for RABL4 based on the alignment of 2 AVAs (BC000566 and AK090708) and a valid accession (NM_006860). The accession numbers for AVAs are followed by their gene description whereas no such description is given for the valid accession, NM_006860. the choice is left to the user to determine whether he wishes to use the sequence extracted based on the identified AVAs or valid accession. A 'View FASTA Sequence' hyperlink is provided for each identified SOE1 position. The 5'-most SOE1 position is always given first on the page followed by the next 5'-most SOE1 position identified. All labels in red were added by the authors for clarity of understanding.

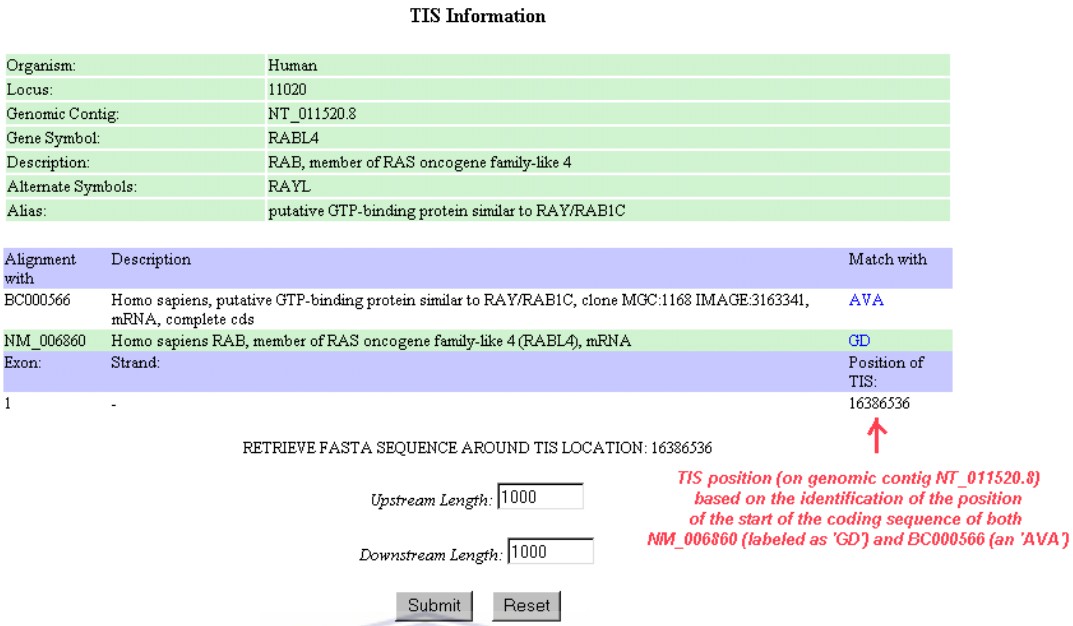


Figure 2.1c:

An example of FIE2's TIS Information page. As can be seen here, the identified TIS position for RABL4 is supported by coding sequences of both NM_006860 (a valid accession, denoted as 'GD' on the page) and BC000566 (an AVA). All labels in red were added by the authors for clarity of understanding.

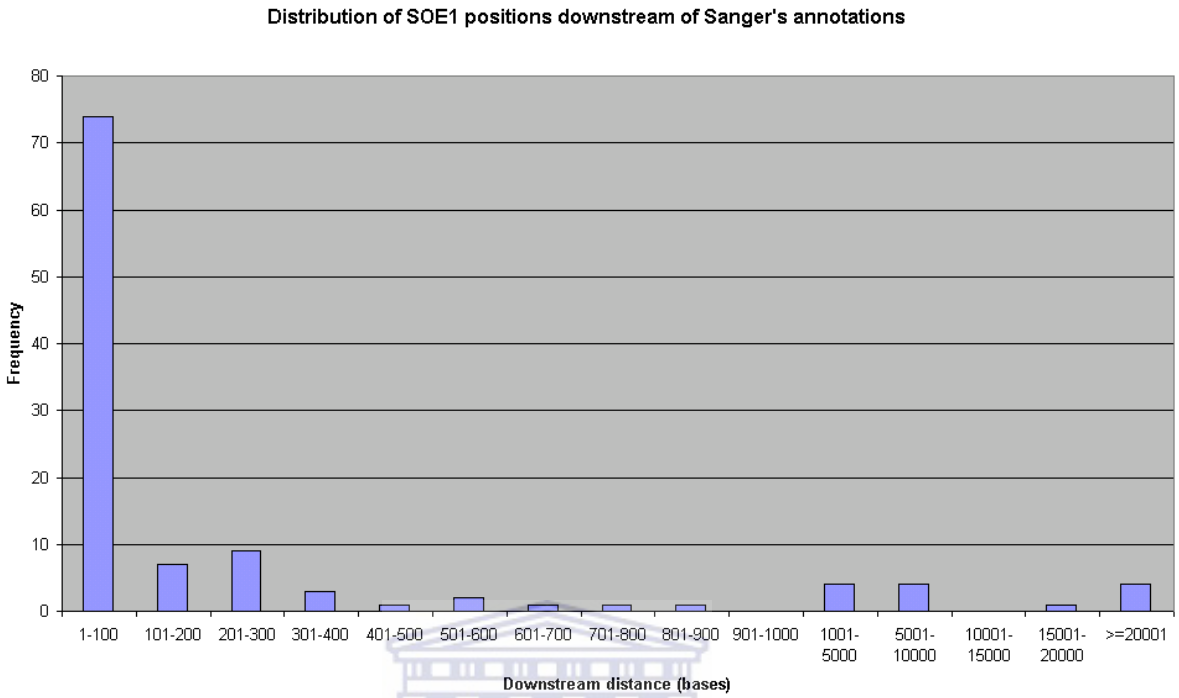


Figure 2.2:

Histogram of the distribution of SOE1 positions extracted by FIE2 (for 112 genes) which were shown to be downstream of Sanger's annotated gene start positions. The "Downstream distance" indicates the difference in distance between Sanger's annotated gene start positions and FIE2's SOE1 position. This distance is measured in terms of the number of bases.

CHAPTER 3: IDENTIFYING KEY COMMON TRANSCRIPTIONAL REGULATORS WITHIN THE PROMOTERS OF THE AMPA RECEPTOR FAMILY OF GENES

3.1 Aim

A multitude of transcription factor (TF) and transcription factor binding site (TFBS) databases and software for the discovery of TFBSs currently exist (Quandt *et al.*, 1995; Kel *et al.*, 1995; Heinemeyer *et al.*, 1999; Kel-Margoulis *et al.*, 2000; Kel-Margoulis *et al.*, 2002; Kel *et al.*, 2003; Bajic *et al.*, 2003; Loots & Ovcharenko, 2004; Sharan *et al.*, 2004; Kim *et al.*, 2005; Guo *et al.*, 2005; Watt *et al.*, 2005; Zhao *et al.*, 2005). With all these bioinformatics resource tools readily available, one might think that finding key TFBSs on a promoter might be as simple as running a BLAST program to find a sequence similarity search. However, this is far from the truth since TFBS search programs like MATCH and MatInspector throw up a fair number of false positive results (Quandt *et al.*, 1995; Kel *et al.*, 2003).

The AMPA receptor family of genes represent the major excitatory amino acid neurotransmitter receptors in the nervous system. However, to date, little work has been carried out to elucidate the key transcriptional elements within the promoters of this family of genes.

The primary aim of this study is to find key transcriptional elements of the AMPA receptor gene family that have been conserved across 3 species (human, rat and mouse) through the use of bioinformatics tools but significantly reducing the number of false positives generated.

3.2 Introduction

Traditionally, DNA footprinting experiments were carried out in the wet lab to elucidate which transcription factor or associated protein binds to the promoters, enhancers or silencers to drive or repress the transcription of genes (Galas & Schmitz, 1978). With the advent of DNA microarray technology, one had high-throughput identification of genomic transcription factor binding sites (TFBSs) *in vivo* through microarray-based readout of chromatin immunoprecipitation assays (so-called 'ChIP-chip') (Ren *et al.*, 2000). And with the availability of next-generation high-throughput sequencing, it is now also possible to map protein-DNA interactions across the genome through the ChIP-Seq method (Johnson *et al.*, 2007; Jothi *et al.*, 2008).

However, these wet lab methodologies for the identification of key transcriptional elements are both time-consuming and expensive. Hence, it would be advantageous to have a computational means of accurately identifying key TFBSs that control a gene's expression. Unfortunately, this is easier said than done. TFBSs are usually short (around 5-15 base-pairs (bp)) and they are frequently degenerate (similar but not identical) sequence motifs. Thus, potential binding sites can occur very frequently in larger genomes such as the human genome. In higher eukaryotes, TFBSs can occur upstream, downstream, or in the introns of the genes that they regulate; in addition, they can be close to or far away from regulated gene(s). Moreover, approximately 95-99% of the human genome does not encode proteins. For all these reasons, it can be very difficult to find TFBSs in noncoding sequences using relatively simple sequence-searching tools like BLASTN. From personal experience, even the use of more sophisticated position weight matrices-based tools, like MATCH (Kel *et al.*, 2003), will generate multiple putative hits for a TFBS along the length of a gene's promoter. It is thus essential to find a way to accurately locate the 'true' TFBSs that control a gene's expression.

Previous work has shown that key transcriptional elements involved in the expression of a gene are over-represented in its promoter (Lin *et al.*, 2004). An analysis of estrogen receptor target genes showed a significant enrichment of putative estrogen response

elements (EREs) in the cis-regulatory regions of these genes. The method of discovering putative TFBSs through a search of over-represented motifs within the promoter has been used repeatedly in bioinformatics (Elkon *et al.*, 2004; Smith *et al.*, 2005).

Here, a novel methodology was used with the aim of identifying key transcriptional regulatory elements and/or transcription factor binding sites that are over-represented and conserved within the promoters of the GRIAs which, in turn, are therefore expected to be essential to the regulation of the expression of GRIA subunits. Using a series of biocomputing procedures described below, both key individual TFBSs and composite elements within GRIA promoters of the human, mouse and rat genes were identified. A composite element is a set of transcriptional regulatory elements and/or TFBSs found in combination on the promoter and usually, in close proximity to each other that works synergistically to control the expression of a gene. An example of this is the IL-4-responsive element in the SOCS-1 promoter which contains three STAT6 and one Ets consensus binding sequences (Travagli *et al.*, 2004). Ets-1 is confirmed to physically interact with STAT6 and IL-4 responsiveness was either partially or totally abolished following specific mutations. Furthermore, exogenous expression of Ets-1 in conjunction with STAT6 activation strongly inhibited expression of a SOCS-1 promoter-luciferase reporter (Travagli *et al.*, 2004).

In addition, this chapter also presents a detailed phylogenetic footprinting study of the human, mouse and rat GRIA1 promoters that reveals consensus sequences within the GRIA1 promoters preserved across the three species and therefore, likely to represent important transcriptional regulatory elements for this gene. Studies with *in situ* hybridization and immunocytochemistry both confirm the widespread expression of GRIA1 in the brain and therefore, substantiates this subunit's importance to AMPA receptor physiology (Boulter *et al.*, 1990; Keinänen *et al.*, 1990; Martin *et al.*, 1993).

The identification of key transcriptional regulatory elements in the GRIA promoters has a greater global significance in that it can be applied in the design of novel gene-targeting constructs. For example, the identification of say, a neuron-specific glutamate receptor

promoter element could possibly be used to deliver future experimental transgene and therapeutic agents to selected neurons in the brain.

3.3 Experimental Procedure

3.3.1 Promoter sequence collection

A total of 10,741 promoter sequences of human genes covering the region -1500 to +1000 (with respect to the transcription start site [TSS]) was collected by the FIE2 program (Chong *et al.*, 2003). The release of the human genomic sequences at the time of collection was NCBI Build 31. The extraction of the human GRIA promoters were done similarly. For the mouse GRIA promoter sequences, the alignment of the individual mouse GRIA genes againsts the then available mouse genomic sequences from the Mouse Genome Sequencing Consortium (MGSCv3) were obtained from NCBI's LocusLink (Pruitt *et al.*, 2000; Maglott *et al.*, 2000; Pruitt & Maglott, 2001). The 5'-end of the extracted mouse GRIA sequences were passed through the Dragon Promoter Finder (without the use of the RepeatMasker) (Bajic *et al.*, 2002) to predict the TSS for these genes. At the start of this study, sequencing of the rat genome had just begun and so, only the rat GRIA1 promoter sequence (AF302117) was used (Borges & Dingledine, 2001). Five TSSs were previously identified for the rat GRIA1 gene (Borges & Dingledine, 2001). For this study, the 5'-most TSS was chosen as a point of reference.

3.3.2 Comparison of target promoters against all other promoters

GRIA promoters (target promoters) were compared with the 10,741 promoters (background promoters) collected by FIE2 in order to determine the over-represented transcriptional elements within GRIA promoters. To do this, all TFBSs from TRANSFAC Professional database ver. 6.2 (Matys *et al.*, 2003) were first mapped to all promoter sequences. This mapping was carried out using the MATCH program (Kel *et al.*, 2003) with the 'minsum' setting. This parameter setting allows for the minimized

sum of false positive (FP) and false negative (FN) predictions of TFBSs. Once that was done, a comparison of the densities of single TFBSs in the target and background promoters was made and the over-representation was calculated, that is, determining how much more dense a particular TFBS is in the GRIA promoters as opposed to the background promoters following the procedure described in Bajic *et al.* (2004). Similarly, the analysis was repeated for all combinations of paired TFBSs where the two TFBSs were no more than 50 nt. apart. The same analysis was also carried out for combinations of three TFBSs, with maximal mutual distance of neighboring TFBSs not greater than 50 nt. Based on such analysis, the top-three ranked single, pair and triplet patterns were selected and these are presented as a model of the GRIA promoters. P-values were calculated using the right-sided Fisher's exact test based on hypergeometric distribution.

3.3.3 Phylogenetic footprinting of GRIA1 promoter region

For this study, I chose to investigate only the promoter region 1000 nt. upstream of the TSS. The human GRIA1 promoter sequence was aligned against the rat GRIA1 promoter using the pairwise local alignment software, Water, from the EMBOSS suite (Rice *et al.*, 2000). Water uses the Smith-Waterman algorithm (Smith & Waterman, 1981) (modified for speed enhancements) to calculate the local alignment. Default parameters for the Water program were used. A similar pairwise alignment was also made with the mouse GRIA1 promoter sequence against the corresponding sequence of the rat GRIA1. The results of the pairwise alignment were studied and conserved regions within the human, mouse and rat GRIA1 promoter region were identified. The sequences for these conserved regions were then passed through the MATCH program (Kel *et al.*, 2003) in order to try and identify the putative TF that might bind these conserved regions. The program setting used was the same as explained above.

3.3.4 Identification of genes with a similar expression pattern to GRIAs

An analysis of human and mouse gene expression data from the Stanford Microarray Database (Gollub *et al.*, 2003) was performed to obtain the top 1% of genes which are

closely coexpressed with GRIAs. This was done by the method previously described by Pellegrino *et al.* (2004). A search was also performed for genes with similar expression patterns using the Human Gene Sorter tool (Kent *et al.*, 2005). The Gene Sorter calculates and displays genes by their similarity in expression to a selected gene. The similarity is calculated as a weighted sum of differences in log expression ratio values using the expression data from such sources as the GNF Gene Expression Atlas 2 (Su *et al.*, 2004). In addition, the NCBI Gene Expression Omnibus' (GEO) "profile neighbor" function was also used to obtain a list of genes that are closely related in expression to the GRIAs, based on data deposited in GEO (Barrett *et al.*, 2003).

3.4 Results

3.4.1 Identification of a unique TFBSs and composite elements combination within the promoters of the GRIA family of genes

Employing the methodology described above, I identified a set of individual TFBSs and composite elements within the GRIA gene family in the region -1500 to +1000 (with respect to the transcription start site [TSS]) that make up a unique promoter profile. For promoter profiling, I selected the set made up by:

1. the top 3 individual TFBSs (which is referred to as 'singles') (see Table 3.1a),
2. the top 3 composite elements each containing a pair of TFBSs that are separated by a distance of no greater than 50 bases (which is referred to as 'pairs') (see Table 3.1b), and,
3. the top 3 composite elements each containing a TFBS triplet, with the distance between each adjacent TFBS being no greater than 50 bases (which is referred to as 'triplets') (see Table 3.1c),

as ranked by over-representation in this analysis.

Appendix 4 gives the coordinates of the 'singles', 'pairs' and 'triplets' found within these 9 GRIA promoters.

3.4.2 Conserved sequences with the promoter of GRIA1 identified through phylogenetic footprinting

Results of the pairwise local alignment software program, Water, from the EMBOSS suite (Rice *et al.*, 2000) for the alignment of the promoter sequences (1000 bps upstream of the TSS) of human GRIA1 and murine GRIA1 against that of rat GRIA1 are given in Figures 3.1a and 3.1b, respectively. The results of the pairwise alignment indicate that rat GRIA1 promoter bears a higher degree of identity to the human GRIA1 promoter (66%) than to the murine GRIA1 promoter (40%) over the aligned regions. This is in contrast to results obtained by an alignment of the genes themselves using BLAST. Simply blasting the rat GRIA1 gene sequence against its human or mouse counterpart with the bl2seq program (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi) shows that these sequences share a degree of identity of 90% over the aligned regions (Tatusova & Madden, 1999). In fact, from UniGene, we can see that the GRIA1 proteins from the 3 species share as much as 99% identity (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene ; Wheeler *et al.*, 2003). Thus, it is evident that there is divergence in the promoter sequences of these genes even though the genes themselves were very much conserved over evolution.

The pairwise alignment of human GRIA1 promoter versus rat GRIA1 promoter and murine GRIA1 promoter versus rat GRIA1 promoter, allowed the identification of sequences which were conserved within the promoters of these genes and thus, are likely to be important to the transcriptional regulation of the GRIA1 gene. Scanning the results of the two pairwise alignment visually, I identified sequences which were conserved across all three promoters. I found, in total, eight conserved regions within the GRIA1 promoters. Using the MATCH program, I attempted to identify the putative TFs that might bind to these conserved regions. Table 3.2 list the sequences and their positions (relative to the TSS) and the putative TFBSs associated with these conserved regions. It

is quite obvious from the results of pairwise alignment (Figure 3.1a and 3.1b) that the relative position of these conserved regions within the GRIA1 promoters is also conserved. This serves to further support the contention that these conserved regions within the promoter might play a significant role in transcriptional regulation.

Of the eight conserved regions, MATCH identified four of these as putative STAT binding sites(1/ human: -944 to -937, rat: -837 to -830, mouse: -760 to -753; 2/ human: -545 to -538, rat: -463 to -456, mouse: -379 to -372; 3/ rat: -690 to -683, mouse: -625 to -618; 4/ rat: -591 to -584, mouse: -484 to -477). However, for two of these putative STAT binding sites, MATCH did not identify the corresponding conserved sequences in the human GRIA1 promoter (-806 to -799 and –706 to -699) as STAT binding sites. There are three plausible explanations for this:

i) the matrix model used by MATCH and thresholds that was selected to identify STAT binding sites is not entirely sensitive (One should note that if the thresholds are relaxed then many more site predictions would be possible. Furthermore, the matrix model of STAT binding sites need not necessarily provide sufficient coverage of real STAT sites);

ii) the conserved sequences in the human, rat and mouse promoters may look very much like a STAT binding site but may, in fact, bind a TF not covered in TRANSFAC (the database of TF matrices used by the MATCH program);

iii) these conserved sequences may not, in fact, bind any TF and therefore, have no significance in transcriptional regulation.

Two other conserved regions were identified as putative HOXA3 (human: -635 to -627, rat: -555 to -547, mouse: -449 to -441) and MAZ (human: -429 to -422, rat: -348 to -341, mouse: -185 to -178) binding sites. The remaining two conserved regions were identified as putative Xvent-1 (rat: -704 to -692, mouse: -641 to -629) and GATA (rat: -658 to -652, mouse: -594 to -585) binding sites in the rat and mouse promoters but MATCH did not recognize the corresponding sequences in the human GRIA1 promoter, -819 to -807 and -774 to -768, as Xvent-1 and GATA binding sites, respectively. It is interesting to note that in the conserved regions of rat: -704 to -683 and mouse: -641 to -618, two putative TF binding sites (Xvent-1 and STAT) are found side-by-side. It is, therefore, very likely

that this region serves as a composite element, possibly for Xvent-1 and STAT to work in concert with each other.

3.4.3 Analysis of genes that are coexpressed with GRIAs

In order to support the claim that the above 47 genes are coregulated / coexpressed with GRIAs, I used available expression data repositories and tools to see if I could find evidence to show that these genes and GRIAs have a statistically correlated coexpression pattern. I used two publicly available resources to find genes that show a similar pattern of expression relative to the GRIA family of genes. The first resource is provided by the UCSC Genome Browser called Gene Sorter (Kent *et al.*, 2005). Mining through the various expression databases available on Gene Sorter, I found 4793 genes that are closely coexpressed with GRIAs. I further discovered that 16 of the 47 genes that I postulated to be coregulated /coexpressed with GRIAs can be found among this list of 4793 genes (Table 3.3).

Next, I used NCBI's (National Center for Biotechnology Information) Gene Expression Omnibus (GEO) to perform a similar search for genes that have a correlated expression profile to that of the GRIAs (Barrett *et al.*, 2005). Here, I found 7354 such genes (termed "profile neighbors" by GEO) expressed in brain, of which 10 are among the list of 47 genes that I postulated to be coregulated /coexpressed with GRIAs (Table 3.3). Of these 10 genes, 6 were the same as those identified by Gene Sorter.

With data provided by Ferdinando DiCunto from the CLEO database (Pellegrino *et al.*, 2004), I analysed the top 1% of genes with a statistically correlated coexpression pattern to that of the GRIAs from the human and mouse gene expression data of the Stanford Microarray Database (Gollub *et al.*, 2003). Here, I found 10 of the 47 genes among the top 1% - 6 genes from the human gene expression data and 4 genes from the mouse gene expression data. Of these 10, 6 were the same as those identified by either Gene Sorter and/or GEO.

Thus, in all, I found 24 genes that are closely coexpressed with GRIAs while also sharing a similar promoter profile, of which, proof of this coexpression for 10 genes are confirmed by 2 or more of the above-mentioned analyses (that is, the analyses of the human and mouse gene expression data, NCBI's GEO and/or the human Gene Sorter program) (Table 3.3). The International Human Genome Sequencing Consortium confirms the existence of 19,599 protein-coding genes. Examining the data for the human gene population, the p-value for enrichment for the 24 genes were calculated to be 4.046736e-002 (after conservative correction for multiplicity testing done with the Bonferroni method) for genes that are closely coexpressed with GRIAs while also sharing a similar promoter profile. The calculation for the p-value for enrichment is derived as follows:

(i) Total number of human genes: 20000 (19,599 protein-coding genes are confirmed by International Human Genome Sequencing Consortium)

(ii) Number of genes coexpressed with four human GRIA genes: ~ 4800 (an average of the number of closely coexpressed genes identified by all three microarray datasets is 4783 genes; that is, [4793 genes (from GeneSorter) + 7354 genes (GEO database) + 2202 genes (Stanford Microarray Database)] divided by 3)

(iii) Number of human genes that have GRIAs' promoter model: 51 (4 GRIA + 47 non-GRIA)

(iv) Number of human genes that have GRIAs' promoter model and are closely coexpressed: 28 (4 GRIA+24 non-GRIA).

If we put these numbers into a contingency table we get

	Genes that coexpress with GRIA	Genes that do not co-express with GRIA	Total
Genes that have promoter model	28	23	51
Genes without promoter model	4772 = (4800-28)	15177 = (15200-23)	19949
Total	4800	15200	20000

From this contingency table, we obtain the p-value for the enrichment of genes with the GRIA promoter model that are also coexpressing with GRIAs, as
p-value = 2.023368e-006

However, the conservative correction for multiplicity testing done with the Bonferroni method gives corrected the p-value of 4.046736e-002, which is below significance threshold of 0.05. Thus, it can be concluded that the enrichment is statistically significant under the given assumptions.

For the UCSC Gene Sorter results, the similarity in expression of two genes is calculated as a weighted sum of differences in log expression ratio values whereas for the analysis of the data from the Stanford Microarray Database and for GEO Profiles' pre-calculated profile neighbors, a calculation of the Pearson's Correlation Coefficient was made to acertain the closeness in expression of two genes. If one were to use the UCSC results as a point of reference, then by two different methods of expression profiling, we can find at least 8 different genes that are not only confirmed to be coexpressed with GRIA but also share a similar promoter profile. These 8 genes (BAI2, IL16, KLK6, CLSTN3, TU3A, KLHL18, BEX1, C20ORF172) therefore represent very strong candidates for further laboratory studies.

3.5 Discussion

3.5.1 Key strengths of the methodology used

Here, I employ the use of MATCH (Kel *et al.*, 2003) in combination with a novel methodology developed by Bajic and co-workers (2004) to find key transcriptional elements that have been conserved across 3 species: human, rat and mouse. This method offers a significant edge over just the conventional use of a TFBS prediction program because

1) it highlights key promoter elements unique to a family of genes by comparing the promoter profile of target gene family against a background of about 10,741 human gene promoters in order to determine the over-represented transcriptional elements within the target promoters, and

2) this method of feature identification is not identifying individual TFBSs or composite elements on a single promoter sequence but a combination of regulatory elements that is uniquely characteristic to several promoter sequences in a gene family. Computational analysis on a single sequence often throws up false positive errors; for example, Borges & Dingledine (2001), using MatInspector, found two putative AP-1 sites within the GRIA1 promoter but electrophoretic mobility shift assays (EMSAs) failed to show binding of c-Jun to these sites. In this case, the GRIA genes from 3 different species were used to identify the unique combination of transcriptional elements present in these genes' promoters and therefore, this could be viewed as an advanced form of phylogenetic footprinting. However, if the gene family was sufficiently big, this technique could perhaps be applied to a single gene family of a single species.

For these reasons, the method used reduces false positive errors normally generated by the conventional TFBS prediction programs and provides a higher level of statistical significance.

Furthermore, it may be possible to use a unique promoter profile to discover novel genes of a gene family, unlike the conventional method of identifying genes of the same family through sequence homology searches. Moreover, as will be illustrated in Chapter 4, this type of promoter profiling also allows us to identify potential co-regulated / co-expressed genes.

3.5.2 Characterization of the GRIA promoter profile

At the start of this study, I speculated that there must be a combination of transcriptional regulatory elements within the promoter of a gene that uniquely controls its expression and therefore, is characteristic or perhaps, even somewhat unique, to that gene and its close family members. If this is so, then this unique combination (set) of transcriptional regulatory elements would most likely be conserved in evolution and identifiable across members of the same gene family in various species of organisms. To prove this, I took the sequence 1500 bases upstream and 1000 bases downstream of the TSS for each gene of the GRIA subfamily of glutamate receptor in human, mouse and rat (for the rat, only the sequence around the TSS of the rat GRIA1 gene was used in this study since the sequencing of the rat genome was at its threshold at the start of this study) and discovered the unique promoter profile consisting of 3 singles, 3 pairs and 3 triplets that characterized well the GRIA promoters (described in Tables 3.1a, 3.1b and 3.1c, respectively). The full MATCH results for each GRIA promoter and the positions of the singles, pairs and triplets (with respect to the TSS) are given in Appendix 4. It was also found that this promoter profile is shared by 47 other gene promoters (discussed further in Chapter 4) out of a field of 10,741 human gene promoters extracted by the FIE2 program (Chong *et al.*, 2003). Although this promoter profile of 3 singles, 3 pairs and 3 triplets is not unique to the GRIA promoters, its uniqueness is still significant as it is represented in less than 0.5 percent of a comprehensive group of gene promoters. I postulate that the reason these 47 gene promoters share this unique promoter profile of 3 singles, 3 pairs and 3 triplets may be because these 47 genes are co-regulated / co-

expressed with the GRIA genes and play an important role in GRIA expression or function. Of these 47 genes, 16 had functions that were yet unknown.

From among the TFBSs and composite elements that I identified to be over-represented on all 9 AMPA receptor gene promoters, I found that there is prior evidence for the existence of the composite element pair, MZF1 and GATA (see Table 1b). Yu and colleagues (2005) found that the solitary ERV-9 long terminal repeat located upstream of the HS5 site in the human beta-globin in erythroid K562 cells contained DNA motifs that bound the ubiquitous factor, NF-Y, and MZF1 and GATA-2. Through protein/protein interactions, NF-Y bound at the CCAAT motif and recruited MZF1 and GATA-2 and stabilized their binding to the neighboring GTGGGGA and GATA motifs.

Results suggest that angiotensin (AngII) activates STAT6 and STAT3 and these transcription factors are involved in the activation of the angiotensinogen (ANG) promoter via their recognition of the St-domain sequence (Mascareno *et al.*, 1998). In addition, a STAT3/Lyf-1/MZF1 composite element located in the promoter region from -238 to -144 of the mouse frizzled-related protein 4 (sFrp4) gene was found to be essential for the promoter activity of sFrp4 (Wong *et al.*, 2003). This suggests that STAT3 and MZF1 may interact with one another and thus, leads us indirectly to believe that STAT6/MZF1/STAT3 may interact and bind to the composite element triplet that was identified on the GRIA promoters (Table 1c).

Cytokine-induced activation of the Fcgamma receptor I promoter required the DNA binding, and the transactivation functions of both Stat1 and PU.1 (Aittomaki *et al.*, 2004). In addition, an analysis of the human CD40 promoter indicates that the two gamma activated sequence (GAS) sites at -521 and -483 and two Ets family member binding sites located at -553 and -447 are important for interferon (IFN)-gamma induction of CD40 transcription (Nguyen & Benveniste, 2000). PU.1/Spi-B binds the distal (-553) while PU.1 binds the proximal (-447) Ets sites. It is unclear how these transcription factors cooperate to switch on CD40 promoter activity but their close proximity might suggest a direct interaction between STAT1/PU.1/Spi-B. Further evidence of STAT1/PU.1

cooperativity can also be seen in IFN-gamma's induction of transcription of macrophage fgl2 gene (Liu *et al.*, 2006). Incidentally, it is believed that IFN-gamma treatment of spinal dorsal horn neurons causes a reduced expression of GluR1 (Vikman *et al.*, 2003). It should also be noted that IFN-gamma can increase the expression of GKLF (Chen *et al.*, 2002). I therefore postulate that the GKLF/PU.1/STAT1 composite element that was identified (Table 1c) may play a role in the IFN-gamma-mediated decrease of GluR1 expression.

3.5.3 Conserved regions within the GRIA1 promoters

Although the human, mouse and rat GRIA1 genes share about a 90% degree of identity, their promoter regions 1000 nt. upstream of the TSS share, on average, only about a 50% degree of identity. However, within the 1000 nt. GRIA1 promoter region, it was possible to distinguish eight distinct regions which were highly conserved across the three species. It is believed that these regions within the promoter are conserved because of their relevance to the transcriptional regulation of GRIA1. What is interesting to note is that the relative positions of these conserved regions within the GRIA1 promoter is also conserved (see Figure 3.1a & 3.1b).

Using the MATCH program (Kel *et al.*, 2003), it was found that at least four of the eight regions to be STAT or STAT-like binding sites (see Table 3). Two other regions were identified by MATCH to be putative HOXA3- and MAZ-binding sites. As for the remaining two conserved regions, they bore similarity to the Xvent-1 and GATA binding sites in the mouse and rat sequences although the results for the corresponding human sequences were not conclusive. In this latter case, it is surprising to note that the conservation of these putative TFBSs is greater in the mouse and rat GRIA1 promoters than in the human GRIA1 promoter. This is despite the fact that, over the aligned regions, the rat promoter sequence bears a higer degree of identity to the human promoter sequence (66%) than it does to the mouse promoter sequence (40.3%).

Previous study indicates that the rat GRIA1 promoter is mostly neuron-specific (Borges & Dingledine, 2001). The neuronal to glial expression ratio of the most neuron-specific constructs was higher than that of GRIA2 due to low GRIA1 promoter activity in glia compared to GRIA2. It was suggested that a glial silencing region exists in the region –391 to –164 (corresponding coordinates as given by Borges & Dingledine (2001) are –686 to –459). Incidentally, I found that one of the eight conserved regions, a putative MAZ-binding site between –348 and -341, is located within this glial silencing region. The remaining 7 conserved regions which include all 4 STAT and/or STAT-like binding sites, the HOXA3 binding site and the Xvent-1-like and GATA-like binding site, were, however, found within the characterized neuron-specific region from –1100 to –448 (corresponding coordinates as given by Borges & Dingledine (2001) are –1395 to –743). With the high concentration of putative STAT and/or STAT-like binding sites in this region of the GRIA1 promoter, the JAK/STAT pathway may play a crucial role in the neuronal specificity of the GRIA1 gene.

Shortening of the 64 bp GA repeat, found between –100 and –37 of the rat GRIA1 promoter, in some constructs reduced the promoter activity in glia while not affecting expression in neurons (Borges & Dingledine, 2001). Although I found a similar 62 bp GA repeat (-191 to -130) in the human GRIA1 promoter, I could find no such repeat sequence in the mouse sequence as far down as 1000 bp downstream of the predicted TSS. Also, a 57 bp region (from -448 to -391) containing an N-box (from -440 to -435) in the rat GRIA1 promoter was found to be a negative regulator of GRIA1 expression (Borges & Dingledine, 2001). Deletion of this region saw an increase in promoter activity in both neurons and glia. Again, I found from the pairwise alignments that although the N-box, is preserved in the human GRIA1 promoter (from -481 to -486), it is not found in the mouse GRIA1 promoter.

3.5.4 Coexpression data supports hypothesis of coregulated genes

By coregulation, I mean that the transcriptional regulation of a particular gene is closely linked to the transcription of GRIA genes due to their shared promoter characteristics.

Coregulated genes are believed to be functionally related and play a significant role in aiding the function and expression of their partner. While coregulation of two genes may result in the coexpression of these two genes in the same tissue, coexpression of two genes in the same tissue does not necessarily mean that they are transcriptionally coregulated since their coexpression may be coincidental. However, we should not also expect two transcriptionally coregulated genes to always be coexpressed since their transcriptional coregulation may take place only under the right physiological conditions; transcription, after all, is affected by various factors, such as, tissue-specificity. Recent studies have shown that if the expression of two or more genes is constantly related throughout many independent microarray datasets, the genes display a significant degree of functional similarity (Lee *et al.*, 2004; Price & Rieffel, 2004). In this context, it is interesting to note that 10 genes were repeatedly shown to be coexpressed with GRIAs by the analysis of the human and mouse data from the Stanford Microarray Database and by the Gene Sorter tool and/or NCBI's GEO. Particularly interesting is the close coexpression of KLK6 with GRIA that was found in human (by Gene Sorter) and in mouse (Stanford Microarray Database) which strongly suggests a functional relationship between these two genes. Phylogenetic conservation has been proposed as a very strong criterion for identifying functionally relevant coexpression links between genes (Stuart *et al.*, 2003). Conservation implies that the coexpression of the gene pairs confers a selective advantage and therefore, these genes are most likely functionally related.

One pertinent question that arises is: what if microarray expression data does not show a coexpression of a particular gene (say, Gene X) with GRIA? Can we DEFINITIVELY say that this is a false positive prediction? The answer is, of course, no. The circumstances under which Gene X would be coregulated with GRIA may not be present during the microarray experiment (for example, perhaps the cells need to be treated with tumor necrosis factor-alpha in order for us to see a coregulated expression of Gene X and GRIA).

In further support of the contention that the 47 genes are coregulated with GRIAs, a search for the expression profiles of all 47 genes in UCSC's Gene Sorter and GEO

confirm that all 47 genes are expressed in the brain (fetal and/or adult) to varying degrees. This is important since AMPA receptors are found predominantly in the nervous system.

Table 3.1a:

The Top 3-ranked individual TFBSs ("Singles") that make up a part of the unique promoter profile of GRIAs (see Section A.5.1 of Appendix 4 for detailed coordinates)

	TFBS	Strand	ORI	p-value
1.	CDP CR1	-ve	1.5499	0.02153
2.	Sp3	+ve	1.5727	0.01888
3.	Bach2	+ve	1.7398	0.00760

Table 3.1b:

The Top 3-ranked composite elements containing pairs of TFBSs ("Pairs") that make up a part of the unique promoter profile of GRIAs (the order in which the TFBS appears in each pair refers to it position within the pair from the 5'→3' direction) (see Section A.5.2 of Appendix 4 for detailed coordinates)

	5'→3'					
	TFBS	Strand	TFBS	Strand	ORI	p-value
1.	GKLF	+ve	PU.1	+ve	2.5702	0.00022632
2.	MZF1	+ve	GATA-2	+ve	2.5796	0.00021901
3.	PU.1	+ve	GKLF	+ve	2.6380	0.00017899

Table 3.1c:

The Top 3-ranked composite elements containing TFBS triplets ("Triplets") that make up a part of the unique promoter profile of GRIAs (the order in which the TFBS appears in each triplet refers to it position within the triplet from the 5'→3' direction) (see Section A.5.3 of Appendix 4 for detailed coordinates)

	5'→3'							
	TFBS	Strand	TFBS	Strand	TFBS	Strand	ORI	p-value
1.	STAT6	-ve	MZF1	+ve	STAT3	-ve	4.3538	2.1958e-006
2.	ELF-1	+ve	STAT1	-ve	Pax-4	-ve	4.4384	1.8460e-006
3.	GKLF	+ve	PU.1	+ve	STAT1	-ve	5.5722	2.3734e-007

Table 3.2:

A table listing the conserved regions within the GRIA1 promoters

Sequence	TFBS	Core	Matrix	Strand
Human -944 TAGAATCC -937 |||||.|. Rat -837 TAGAACCA -830 |.|||.|| Mouse -760 TGGAAGCA -753	STAT5A	0.998	0.989	
	STAT5A	0.998	0.988	-ve
	STAT3	0.994	0.993	
Human -635 ATTTCTAGG -627 ||||||||| Rat -555 ATTTCTAGG -547 .....|||| Mouse -449 GAAATTAGG -441	HOXA3	0.971	0.836	
	HOXA3	0.971	0.836	-ve
	HOXA3	1.000	0.955	
Human -545 TAGAAAAA- -538 .|||||.| Rat -463 GAGAAAGT- -456 ..||| || Mouse -379 AGGAA-GTA -372	STAT5A	1.000	0.997	
	STAT5A	1.000	0.991	-ve
	STAT5A	0.977	0.976	
Human -429 GGGGAGCG -422 ||||||.| Rat -348 GGGGAGGG -341 .|||||| Mouse -185 AGGGAGGG -178	MAZ	1.000	0.943	
	MAZ	1.000	1.000	+ve
	MAZ	1.000	0.970	
Human -819 TTTTAAATGCCTA -807 .||.||||.|||| Rat -704 GTTCAAATACCTA -692 .|.||||....| Mouse -641 ATACAAATTGGAA -629	No match			
	Xvent-1	1.000	0.862	-ve
	Xvent-1	1.000	0.897	
Human -806 GGATTGAA -799 ||.||.|| Rat -690 GGTTTCAA -683 .|..... Mouse -625 AGGAAGTC -618	No match			
	STAT4 / STAT1	0.988 / 0.960	0.986 / 0.955	+ve
	STAT6 / STAT3 / STAT5A / STAT4 / STAT1	1.000 / 0.994 / 0.977 / 0.967 / 0.957	1.000 / 0.994 / 0.977 / 0.969 / 0.956	
Human -774 --AGGTAGA- -768 ||.|||| Rat -658 --AGATAGA- -652 .|||||. Mouse -594 TATGATAGTA -585	No match			
	GATA	0.984	0.979	+ve
	GATA-2 / GATA-3 / GATA-6	1.000 / 1.000 / 0.989	0.988 / 0.976 / 0.986	

Human -706 TCTATCTG -699 .||.|||. Rat -591 CCTTTCTC -584 |||||||| Mouse -484 CCTTTCTC -477	No match			+ve
	STAT5A / STAT4 / STAT1	1.000 / 0.988 / 0.971	0.992 / 0.971 / 0.967	
	STAT5A / STAT4 / STAT1	1.000 / 0.988 / 0.971	0.992 / 0.971 / 0.967	

Table 3.3:

List of genes that are closely coexpressed with GRIAs as determined by analyses of independent gene expression datasets.

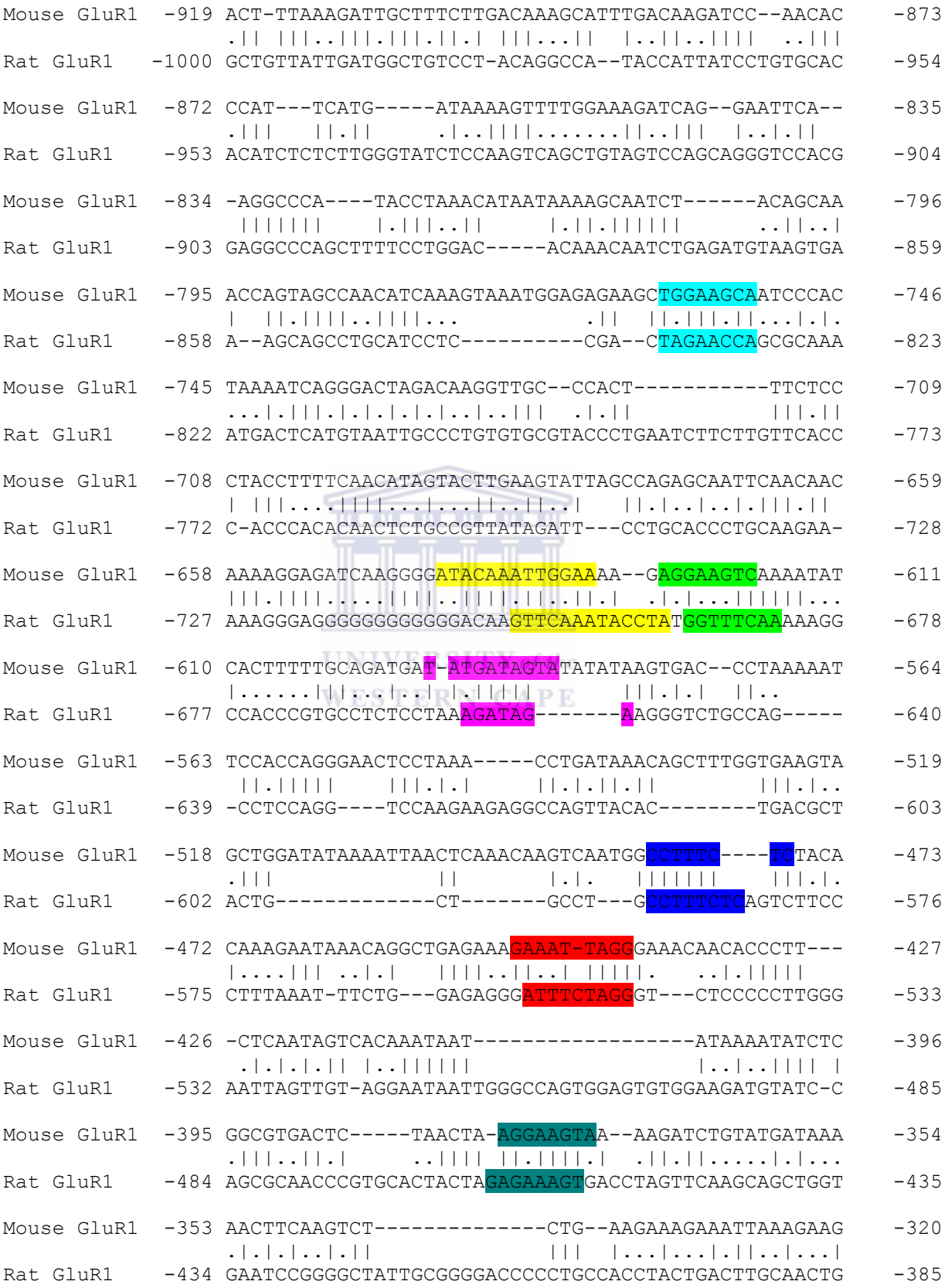
	Human Gene ID (Mouse Gene ID given in brackets)	Gene Symbol	Stanford Microarray Database (Human)	Stanford Microarray Database (Mouse)	UCSC Human Gene Sorter	NCBI GEO
1	576	BAI2	√		√	√
2	911	CD1C	√			
3	2323 (14256)	FLT3LG		√		
4	3603	IL16			√	√
5	4793	NFKBIB	√			
6	5279	PIGC	√			√
7	5653 (19144)	KLK6		√	√	
8	5865	RAB3B			√	
9	8674	VAMP4				√
10	9746	CLSTN3			√	√
11	11131	CAPN11			√	
12	11170	TU3A	√		√	√
13	23276	KLHL18			√	√
14	25852	ARMC8				√
15	27120 (50722)	DKKL1		√		√
16	54093	C21ORF18			√	
17	54897	FLJ20321			√	
18	55244 (67473)	FLJ10847		√		
19	55859	BEX1			√	√
20	56999	ADAMTS9			√	
21	64225	ARL6IP2			√	
22	64577	ALDH8A1			√	
23	79980	C20ORF172	√		√	
24	80227	WDR71			√	

```
Human GluR1  -1000 TACACCAG-----AC--ATGACCAGC------------------ATC-CA    -977
                   |.||.|||     ||  |.|.|||||                   ||| .|
Rat GluR1     -919 TCCAGCAGGGTCCACGGAGGCCCAGCTTTTCCTGGACACAAACAATCTGA    -870

Human GluR1   -976 GA---AAGACCTGAAAGAAGCTTGAATCCTCTCACTAGAATCCCTGCAAA    -930
                   ||   |||   ||||||.|||.||.||||||..|||||||.|...|||||
Rat GluR1     -869 GATGTAAG---TGAAAGCAGCCTGCATCCTCCGACTAGAACCAGCGCAAA    -823

Human GluR1   -929 ATGACTCATGTAATTGCTCTGTGTAAGTATCCTTAGTCTTTATTGT----    -884
                   ||||||||||||||||.|||||..|||.|||.|.|| ||.||||
Rat GluR1     -822 ATGACTCATGTAATTGCCCTGTGTGCGTACCCTGAATC-TTCTTGTTCAC    -774

Human GluR1   -883 ACACCCACACGATTCTGATGCTATAGACTCCTG-----TGGAATGCAGGG    -839
                   .|||||||||.|.||||..|.||||||.|||||     ||.||      |
Rat GluR1     -773 CCACCCACACAACTCTGCCGTTATAGATTCCTGCACCCTGCAA------G    -730

Human GluR1   -838 AAAGAGAGA------AGGGGGCCCATTTTAAATGCCTA-GGATTGAAAAG    -796
                   |||.||.||      .|||||.|.|.||.||||.|||| ||.||.||||
Rat GluR1     -729 AAAAAGGGAGGGGGGGGGGGGACAAGTTCAAATACCTATGGTTTCAAAA-    -681

Human GluR1   -795 AGACCA-CCGTTTCACTTGTAAAGGTAGACAGGGACTGTCAAATACCTGG    -747
                   ||.||| ||||..|.|||..||||||.|||| ||||.|||.||....||.||
Rat GluR1     -680 AGGCCACCCGTGCCTCTCCTAAAGATAGA-AGGGTCTGCCAGCCTCCAGG    -632

Human GluR1   -746 TCAAAATACCTGCCAGT--CACTCCAGATCCTCCCTTGTTTGTCTATCTG    -699
                   ||.||..|...|||||  ||||...|.|.||.|  ||..||.||.||
Rat GluR1     -631 TCCAAGAAGAGGCCAGTTACACTGACGCTACTGC--TGCCTGCCTTTC--    -586

Human GluR1   -698 TCATTCCTTCCATTAGGAGAGAGAAAGCTTTTTTTTTTTTTTCCTTTAAA    -649
                   |||.||                              ||.||||||||
Rat GluR1     -585 TCAGTC---------------------------------TTCCCTTTAAA    -569

Human GluR1   -648 TTTCCTAG-GAGGGATTTCTAGGGTCTTTCCCTCAGGAATTAGTTGTAGG    -600
                   ||| ||.| |||||||||||||||..||||..|||||||||||||||
Rat GluR1     -568 TTT-CTGGAGAGGGATTTCTAGGGTCTCCCCCTTGGGAATTAGTTGTAGG    -520

Human GluR1   -599 AATAATTGGGCCAGTGGAGTGCAGGAGATATATCCAGCGCAGCCCATGCA    -550
                   ||||||||||||||||||..|.||||.||||||||||.|||.||||
Rat GluR1     -519 AATAATTGGGCCAGTGGAGTGTGGAAGATGTATCCAGCGCAACCCGTGCA    -470

Human GluR1   -549 CTCCTAGAAAAAGTGACCTAGATCAAGCAGCTGGTGGATTGAGGACTATT    -500
                   ||.|||||.||||||||||||.||||||||||||||.||...||.|||||
Rat GluR1     -469 CTACTAGAGAAAGTGACCTAGTTCAAGCAGCTGGTGAATCCGGGGCTATT    -420

Human GluR1   -499 GTGGGGACCCCCTGCCACCTACTGACTTACAGCTGAACCCACATTCCC--    -452
                   |.|||||||||||||||||||||||.||.||||||||.|||||||
Rat GluR1     -419 GCGGGGACCCCCTGCCACCTACTGACTTGCAACTGAACCCGCATTCCCAA    -370

Human GluR1   -451 -AGCAGCTT---CAGCCTGGGGGCTGGGGGAGCGGGC------AGACCGA    -412
                    ||||.|||   |  |||   |||.|||||| ||||      |||   |
Rat GluR1     -369 TAGCATCTTATCC--CCT---GGCAGGGGGAG-GGGCTGGGAGAGA---A    -329

Human GluR1   -411 GCTCAGAAAGGCAGGGGAGGGTAAAGAGGACTGTGGGGTT-GCCCCTTTC    -363
                   ||..||.||||..|||||  .||.|.|.||||.|||||.|| .|||.|||.
Rat GluR1     -328 GCAGAGGAAGGTGGGGGA-CGTGAGGGGGACAGTGGGTTTCTCCCTTTTG    -280
```

```
Human GluR1   -362 AGGACCAAGTGCCACGTGTCACACACCC-CCACC---------TCCACCT    -323
                   .|||  |||              ||||| |||||         |||.||.
Rat GluR1     -279 GGGA--AAG--------------CACCCTCCACCTTCCATCCTTCCTCCC    -246

Human GluR1   -322 TT-------CTGCACACACA----GAAAGGAGGATAAGGTGAGGATGG-G    -285
                   ||       ||||||.||||    ||||..||  |.||||||||||| |
Rat GluR1     -245 TTCCATCCCCTGCACCCACAGGGGGAAAAAAG--TGAGGTGAGGATGGAG    -198

Human GluR1   -284 AGGAAGGGG-GAACAGGTAGGGAGGTC--GGCTGTGGAACTCC--AAGCT    -240
                   |||.||||| ||..||.|.||||||.|  .||  .||||.|||  .||||
Rat GluR1     -197 AGGCAGGGGTGATGAGTTGGGGAGGCCAATGC--AGGAATTCCGAGAGCT    -150

Human GluR1   -239 AGCTCGGTGGGTATTAGCAT--AGAGCTTGCTGCCTGTGTGAGTGTGAGG    -192
                   .|||| .|||||||.|||||  ||||||.||.|||||||.||||||||||
Rat GluR1     -149 GGCTC-CTGGGTATCAGCATAGAGAGCTGGCAGCCTGTGAGAGTGTGAGG    -101

Human GluR1   -191 GGGAGAGCGAGAGAGAGCAAGGGAGGGAGAGAGAGGCAGGCTGCGAG-GG    -143
                   |.|||||.|||||||||..||.|||.|||||||||..||...|.||| |.
Rat GluR1     -100 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA    -51

Human GluR1   -142 GAGAG-GAGAGGGA-----------GTG---GGGGAGCCAGCGCT-CCAG    -109
                   ||||| |||||.||           |||   |.||||..|| ||| ||.|
Rat GluR1      -50 GAGAGAGAGAGAGAAACACGGGAGGGTGAGAGAGGAGAGAG-GCTGCCTG    -2

Human GluR1   -108 CT   -107
                   ||
Rat GluR1       -1 CT   0
```

Figure 3.1a

Pairwise alignment of Human GRIA1 against Rat GRIA1 promoter sequences (1000 bps upstream of the TSS). Phylogenetic footprinting obtained by the alignment of these two promoter sequences shows the position of highly conserved sequences that correspond to the TFBS identified in Table 3.2 above.


Mouse GluR1 -919 ACT-TTAAAGATTGCTTTCTTGACAAAGCATTTGACAAGATCC--AACAC -873
Rat GluR1 -1000 GCTGTTATTGATGGCTGTCCT-ACAGGCCA--TACCATTATCCTGTGCAC -954
Mouse GluR1 -872 CCAT---TCATG-----ATAAAAGTTTTGGAAAGATCAG--GAATTCA-- -835
Rat GluR1 -953 ACATCTCTCTTGGGTATCTCCAAGTCAGCTGTAGTCCAGCAGGGTCCACG -904
Mouse GluR1 -834 -AGGCCCA----TACCTAAACATAATAAAAGCAATCT------ACAGCAA -796
Rat GluR1 -903 GAGGCCCAGCTTTTCCTGGAC-----ACAAACAATCTGAGATGTAAGTGA -859
Mouse GluR1 -795 ACCAGTAGCCAACATCAAAGTAAATGGAGAGAAGCTGGAAGCAATCCCAC -746
Rat GluR1 -858 A--AGCAGCCTGCATCCTC----------CGA--CTAGAACCAGCGCAAA -823
Mouse GluR1 -745 TAAAATCAGGGACTAGACAAGGTTGC--CCACT----------TTCTCC -709
Rat GluR1 -822 ATGACTCATGTAATTGCCCTGTGTGCGTACCCTGAATCTTCTTGTTCACC -773
Mouse GluR1 -708 CTACCTTTTCAACATAGTACTTGAAGTATTAGCCAGAGCAATTCAACAAC -659
Rat GluR1 -772 C-ACCCACACAACTCTGCCGTTATAGATT---CCTGCACCCTGCAAGAA- -728
Mouse GluR1 -658 AAAAGGAGATCAAGGGGATACAAATTGGAAAA--GAGGAAGTCAAAATAT -611
Rat GluR1 -727 AAAGGGAGGGGGGGGGGGGACAAGTTCAAATACCTATGGTTTCAAAAAGG -678
Mouse GluR1 -610 CACTTTTTGCAGATGAT-ATGATAGTATATATAAGTGAC--CCTAAAAAT -564
Rat GluR1 -677 CCACCCGTGCCTCTCCTAAAGATAG-------AAGGGTCTGCCAG----- -640
Mouse GluR1 -563 TCCACCAGGGAACTCCTAAA-----CCTGATAAACAGCTTTGGTGAAGTA -519
Rat GluR1 -639 -CCTCCAGG----TCCAAGAAGAGGCCAGTTACAC--------TGACGCT -603
Mouse GluR1 -518 GCTGGATATAAAATTAACTCAAACAAGTCAATGGCCTTTC----TCTACA -473
Rat GluR1 -602 ACTG----------CT-------GCCT---GCCTTTCTCAGTCTTCC -576
Mouse GluR1 -472 CAAAGAATAAACAGGCTGAGAAAGAAAT-TAGGGAAACAACACCCTT--- -427
Rat GluR1 -575 CTTTAAAT-TTCTG---GAGAGGGATTTCTAGGGT---CTCCCCCTTGGG -533
Mouse GluR1 -426 -CTCAATAGTCACAAATAAT----------ATAAAATATCTC -396
Rat GluR1 -532 AATTAGTTGT-AGGAATAATTGGGCCAGTGGAGTGTGGAAGATGTATC-C -485
Mouse GluR1 -395 GGCGTGACTC-----TAACTA-AGGAAGTAA--AAGATCTGTATGATAAA -354
Rat GluR1 -484 AGCGCAACCCGTGCACTACTAGAGAAAGTGACCTAGTTCAAGCAGCTGGT -435
Mouse GluR1 -353 AACTTCAAGTCT----------CTG--AAGAAAGAAATTAAAGAAG -320
Rat GluR1 -434 GAATCCGGGGCTATTGCGGGGACCCCCTGCCACCTACTGACTTGCAACTG -385

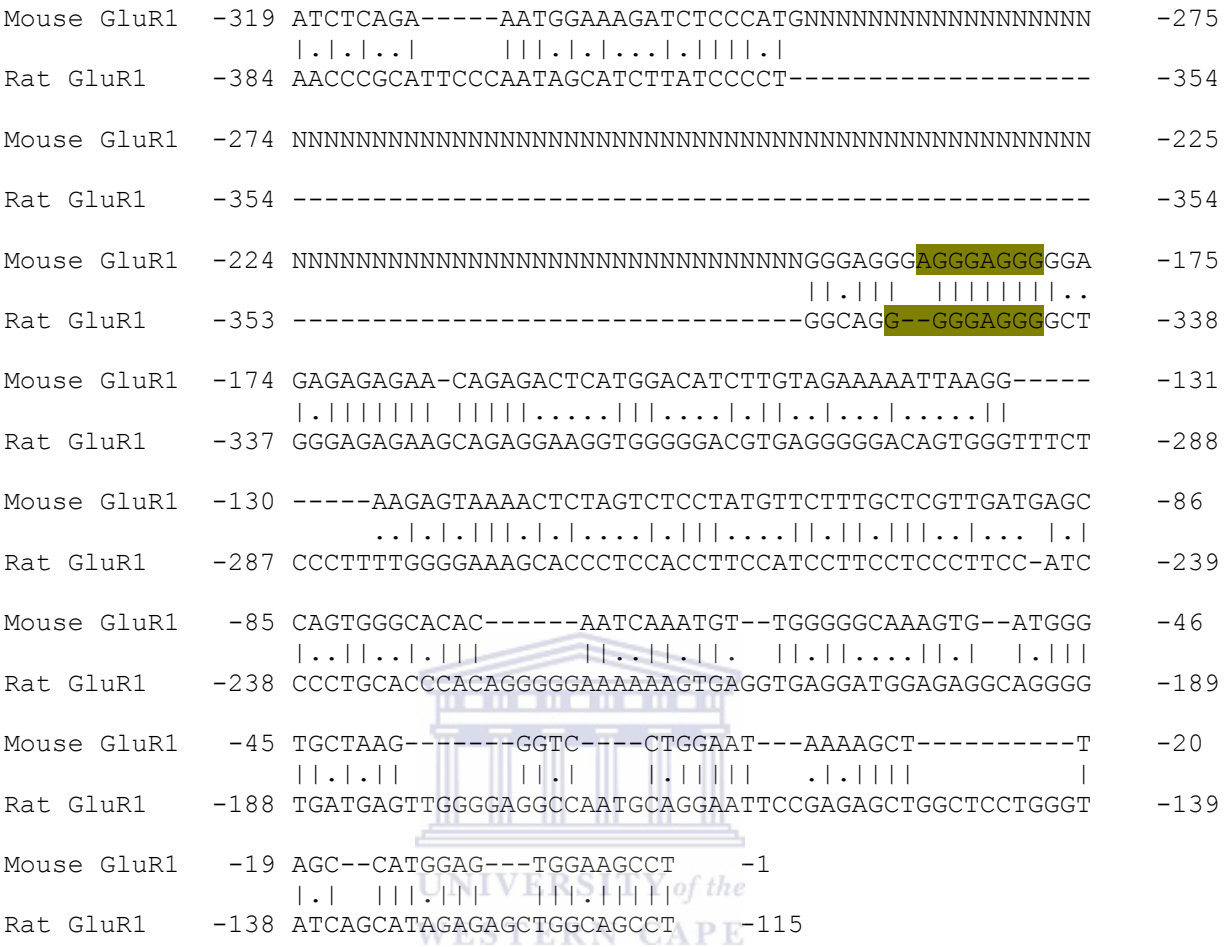


Figure 3.1b

Pairwise alignment of Murine GRIA1 against Rat GRIA1 promoter sequences (1000 bps upstream of the TSS). Phylogenetic footprinting obtained by the alignment of these two promoter sequences shows the position of highly conserved sequences that correspond to the TFBS identified in Table 3.2 above.

CHAPTER 4: AN ANALYSIS OF GENES THAT ARE IDENTIFIED AS BEING CO-REGULATED/CO-EXPRESSED WITH THE GRIA GENE FAMILY

4.1 Aims

As mentioned in Section 3.5.2 in the last chapter, by scanning through the 10,741 background promoter sequences, it was found that the promoter profile identified for the GRIA gene family was shared by 47 other gene promoters. It is believed that the reason these 47 gene promoters share this unique promoter profile of 3 singles, 3 pairs and 3 triplets with the GRIA promoters is due to the fact that they are co-regulated / co-expressed with the GRIA genes and play an important role in GRIA expression or function. Of these 47 genes, 16 had functions that were yet unknown. The remaining 31 genes code for proteins that have physiological and cellular roles that include trafficking of receptors to the cell surface, maintaining or changing cellular morphology, cell growth, transcription, protein biosynthesis and breakdown, signal transduction and even, in retinoic acid metabolism.

It is the aim of this chapter to list these 47 genes and to present supporting evidence to show how seven of these genes could be linked to GRIA expression or function.

4.2 Introduction

There are many attempts in bioinformatics to record, chart and/or predict protein-protein interactions. Several web resources offer users a look at experimentally verified protein-protein interactions:

1) the Database of Interacting Proteins (DIP) is a database of protein-protein interactions but it uses both manual curation by experts and automated text-mining of published literature to identify these interactions (Xenarios *et al.*, 2000; Marcotte *et al.*, 2001;

Salwinski *et al.*, 2004). Each resulting DIP entry reports information about the two interacting proteins, the protein domains and range of amino acids involved, the curator, date of entry and updating and the articles describing the interaction, and the corresponding experiments. For example, a search on a single protein returns all of the interactions recorded in DIP in which that protein participates. (http://dip.doe-mbi.ucla.edu/).

2) the Biomolecular Interaction Database (BIND) is another curated database containing information pertaining to molecular interactions, molecular complexes and pathways of interactions mediating cellular functions (Bader *et al.*, 2003; Alfarano *et al.*, 2005). Like DIP, it also uses a computational text-mining approach to identify relevant published literature describing protein-protein interactions (Donaldson *et al.*, 2003). These literature is then passed through human review before entry into BIND. (http://bind.ca)
3) the Kyoto Encyclopedia of Genes and Genomes (KEGG) is also freely available web resource that among other things offers a "Pathway" database that contains graphical representations of cellular processes, such as metabolism, membrane transport, signal transduction and cell cycle (Kanehisa & Goto, 2000; Kanehisa *et al.*, 2004) (http://www. genome.ad.jp/kegg/).

Several prediction methods have also been utilized to identify putative protein-protein interactions, such as the structure-based multimeric threading method and the domain fusion method (Marcotte *et al.*, 1999; Lu *et al.*, 2002; Lu *et al.*, 2003; Ng *et al.*, 2003).

In all these software and databases, we see that the emphasis is on first-order interactions, that is, direct interactions between two proteins. However, it should be recognized there are also indirect interactions that may aid in or is essential for the regulation of a protein's function or expression. For example, a protein which curbs the over-expression of another protein by interfering with its transcriptional machinery.

In the course of this work, I discovered that it may be possible to identify both direct and indirect interactions, not through a study of the protein itself, but through a study of the

transcriptional elements in the gene's promoter. It is natural to assume that when a protein is expressed, the cell's transcriptional machinery would also transcribe the genes coding for other accessory proteins which are necessary for the former's functioning. In order for this to happen, the promoters of the target protein and it's accessory proteins would need to share certain common transcriptional regulatory elements. Therefore, if the key transcriptional elements within the promoter of our target protein were known, then we could identify interacting partners (whether direct or indirect) by searching for other proteins with the same key transcriptional elements as our target protein.

In this chapter, I describe the identification of 47 genes which are believed to be co-regulated and/or co-expressed with the GRIA family of genes and how they may interact with and, aid in or control the functioning of AMPA glutamate receptors.

4.3 Method & Results

As described in Chapter 3, all TFBSs from TRANSFAC Professional database ver. 6.2 (Matys *et al.*, 2003) were first mapped to all 10,741 background promoter sequences. This mapping was carried out using the MATCH program (Kel *et al.*, 2003) with the 'minsum' setting. This parameter setting allows for the minimized sum of false positive (FP) and false negative (FN) predictions of TFBSs. Scanning through the promoters of the 10,741 genes extracted by FIE2, it was found that the unique promoter profile of 3 singles, 3 pairs and 3 triplets described above (Section 3.4.1; Tables 3.1a, b & c) are shared by 47 other genes. These 47 genes are listed and categorized by function in Table 4.1. Of these 47 genes, 16 genes code for proteins of unknown function. The remaining 31 genes code for proteins that have physiological and cellular roles that include trafficking of receptors to the cell surface, maintaining or changing cellular morphology, cell growth, transcription, protein biosynthesis and breakdown, signal transduction and even, in retinoic acid metabolism.

4.4 Discussion

Sixteen of the 47 genes whose promoter profile contained the same 3 singles, 3 pairs and 3 triplets as the GRIA genes were genes with unknown function. However, I categorized the remaining 31 genes into 14 functional groups whose functions ranged from cellular trafficking to transcriptional regulation (see Table 4.1 for more details). Of these 31 genes, I chose to study, at random, 7 genes (VAMP4, Rab3B, FKBP8, 3-OST-3_A, CLSTN3, SOCS1 and IκBβ) in detail and to provide supporting evidence for their involvement in AMPA receptor function. These 7 genes fall into 3 functional categories, namely, cellular trafficking, cellular morphology and structure, and transcriptional regulation.

4.4.1 Transcriptional regulation

4.4.1.1 Regulation of GRIA expression

Biological systems are filled with feedback systems. I believe that the expression of the GRIA genes should be no different and there should be checks and balances to regulate its expression. One of the primary means of regulating GRIA expression would be at the transcriptional level.

Below, I look at 2 genes (SOCS1 and IκBβ) which I believe are co-regulated / co-expressed along with the GRIA genes so as to regulate the latter's transcription.

4.4.1.2 Suppressor of cytokine signaling 1 (also known as JAK binding protein; STAT-induced STAT inhibitor-1; cytokine-inducible SH2 protein 1; Tec-interacting protein 3) (SOCS1)

SOCS1 belong to a family of proteins that are functionally related by their ability to negatively regulate cytokine and growth factor signaling (Starr *et al.*, 1997). The SOCS family of proteins inhibit signaling by either 1) an inhibition of JAK kinase activity or 2) binding with the activated cytokine receptor (Cooney, 2002).

The CNS expresses an array of cytokines and chemokines (Nitta, 1998; Benveniste, 1998; Asensio & Campbell, 1999). Thus, it would not be surprising that there are mechanisms in place to regulate the activity of cytokines and chemokines. To this end, Polizzotto and coworkers (2000) have also shown the specific expression of regulatory SOCS genes in both developing and mature mouse.

Functional relationship between AMPA receptors, and cytokines and chemokines.

It has been shown that GRIAs are co-expressed with the chemokine receptor CXCR2 and that these two receptors form a multi-protein complex which negatively modulates CXCL2-induced cerebellar granule neuron (CGN) migration (Limatola *et al.*, 2003). The involvement of GRIAs on CGN migration is suggested by the absence of migration in postnatal day 7 (p7) neurons that express high levels of GRIAs but instead were found to migrate upon treatment with GRIA antagonist, CNQX (Limatola *et al.*, 2003). In HEK cells too, co-expression of GRIA1 with CXCR2 greatly impairs CXCL2-induced chemotaxis (Limatola *et al.*, 2003). In addition, immunoprecipitation shows CXCR2 to be associated with GRIAs in p7 CGN and with GRIA1 co-expressed in HEK cells (Limatola *et al.*, 2003). The data, thus, suggest a direct coupling between GRIAs and CXCR2. This interaction is specific since, for example, CXCR4-mediated chemotaxis is not affected by CNQX (Limatola *et al.*, 2003). Just as GRIAs affect CXCR2-mediated chemotaxis, it would seem that CXCR2 also has an effect on GRIAs' properties. GRIA1 coexpressed with CXCR2 have their glutamate dose-response curve shifted to the left (Lax *et al.*, 2002). Furthermore, CXCL2 stimulation of CXCR2 significantly enhances the amplitude of AMPA-type glutamatergic spontaneous EPSCs as a result of increased binding site cooperativity (Lax *et al.*, 2002). Thus, coupling of CXCR2 with GRIAs may modulate the functional profile of the GRIAs.

CXCR2 is a metabotropic G-protein coupled receptor (GPCR) (Thomas *et al.*, 1991a). Although GRIAs are traditionally accepted as ligand-gated ion channels, recent studies have indicated that they may have metabotropic-like properties. GRIA receptor stimulation has been found to (i) activate extracellular signal-regulated kinases (ERKs) in

a phosphatidylinositol 3-kinase (PI3K)-dependent manner in striatal neurons, (ii) trigger mitogen-activated protein kinase (MAPK) activation in cortical neurons via a novel mechanism in which G-protein beta gamma dimers bind to a Ras protein complex causing the activation of Ras, Raf kinase, MEK-1, and finally ERK, (iii) mediate an inhibition of adenylate cyclase activity in cortical neurons via the activation of a G_i protein which is independent of Ca^{2+} and Na^{+}, as a result of an association through the GRIA1 subunit, and (iv) activates the tyrosine kinase Lyn in CGNs (Wang & Durkin, 1995; Wang *et al.*, 1997; Hayashi *et al.*, 1999; Perkinton *et al.*, 1999). In all cases stated above, currents generated by channel opening did not seem to mediate these observed AMPA-evoked metabotropic-like properties.

Thus, it was not surprising to find that in CGNs, treatment with CXCR2 ligands, interleukin-8 (IL-8) and growth-related gene product β (GROβ), and AMPA induced the activation of PI3K and of ERK pathways (Limatola *et al.*, 2002). Treatment either with ERK kinase (MEK) inhibitor, PD98059, or with CNQX abolished AMPA-mediated neurotrophic activity while PI3K inhibitors, LY294002 and wortmannin, blocked GROβ/IL-8 –mediated neurotrophic activity (Limatola *et al.*, 2002). Therefore, AMPA-mediated neurotrophic activity acts via the ERK pathway while chemokine (GROβ/IL-8)-mediated neurotrophic activity acts via the PI3K signalling pathway.

CXCR2 activation results in SOCS1 expression

It would seem that some, if not all, chemokine GPCRs undergo receptor dimerization, as a result of receptor activation, to trigger a signalling cascade through the JAK/STAT pathway (Mellado *et al.*, 1998; Rodriguez-Frade *et al.*, 1999a, b; Vila-Coro *et al.*, 1999; Soriano *et al.*, 2003). CXCR2 is no different and when activated by its ligand, macrophage inflamatory protein-2 (MIP-2), it also passes its signal down the JAK/STAT pathway since an upregulation of STAT3 can be observed in MIP-2-treated mice after hepatic resection (Ren *et al.*, 2003). Furthermore, inhibition of CXCR2 with anti-CXCR2 resulted in a decrease in STAT3 levels (and consequently, decrease in baseline hepatocyte proliferation) (Ren *et al.*, 2003). Coincidentally, the promoter of the SOCS1

gene contains putative STAT3 and STAT6 binding sites as well as a potential gamma activated sequence (GAS) binding site for the STAT1 homodimer (Krebs & Hilton, 2000). The presence of these binding sites therefore support the possible induction of the SOCS1 gene by STAT3, STAT6 and STAT1. Consistent with this finding was the observation that transfection of a dominant-negative mutant of STAT3 blocked the IL-6 or leukemia-inhibitory factor (LIF) induction of SOCS1 mRNA expression in murine myeloid leukemia cells (M1 cells), indicating that STAT3 can, in fact, stimulate the expression of SOCS1 (Naka *et al.*,1997). SOCS1 can directly associate with high affinity with all four types of JAKs and directly inhibit their catalytic activity and thus, not surprisingly, inhibit STAT3 activation (Naka *et al.*, 1997; Endo *et al.*, 1997; Yasukawa *et al.*, 1999; Nicholson *et al.*, 1999; Yasukawa *et al.*, 2003).

Thus, there is evidence to show that the interaction between the chemokine receptor, CXCR2, and AMPA receptors and the subsequent activation of the JAK/STAT pathway with the eventual upregulation of SOCS1 mediated by, possibly, STAT3. We believe that the common promoter elements shared between the SOCS1 and GRIAs, might allow SOCS1 to be co-regulated (not co-expressed) with GRIAs. That is to say, the transcription of SOCS1 might require additional transcription factors (for example, STAT3) to those held in common with GRIA gene expression and therefore, occur at a later stage. This staggered expression of GRIAs and SOCS1 might serve as a feedback mechanism to check the actions of AMPA receptors.

Implication of SOCS1 expression on GRIA gene expression: STAT binding sites within GRIA gene promoters

All 3 triplets of the unique promoter profile (Table 3.1c) contain at least one putative STAT binding site indicating that STAT play an important regulatory role in the expression of GRIA genes. Furthermore, the phylogenetic footprinting study with GRIA1 also found that at least four (out of the eight) conserved regions to be STAT or STAT-like binding sites (Table 3.2). This further supports the suggestion that SOCS1, a

STAT-induced STAT inhibitor, might be co-expressed in the same cell as GRIAs and in turn, act to prevent the overexpression of the GRIA genes.

4.4.1.3 Inhibitor of nuclear factor of kappa light polypeptide gene enhancer in B-cells, beta (IκBβ)

The IκBs and their relationship with the transcription factor, NF-κB.

The transcription factor, nuclear factor-κB (NF-κB), has been implicated in immune and inflammatory responses, cell proliferation and apoptosis (Li & Verma, 2002; Burke, 2003; Gaur & Aggarwal, 2003; Kucharczak *et al.*, 2003). NF-κB exists in quiescent cells in a dormant state in the cytoplasm through their stable association with IκB inhibitor proteins such as, IκBα and IκBβ (Li & Verma, 2002). IκBα and IκBβ bind predominantly to NF-κB p65/p50 heterodimers *in vivo* (Ganchi *et al.*, 1992; Thompson *et al.*, 1995). Activation of NF-κB is mediated through the IκB kinase (IKK) complex, which functions to phosphorylate two serine residues: Ser^{32} and Ser^{36} of IκBα and Ser^{19} and Ser^{23} of IκBβ (Regnier *et al.*, 1997; Mercurio *et al.*, 1997; DiDonato *et al.*, 1997). Phosphorylation of these residues causes the ubiquitination and subsequent degradation of IκB proteins by the 26S proteasome complex (Karin & Ben-Neriah, 2000). Upon loss of IκB, NF-κB is free to translocate to the nucleus where it binds its cognate DNA sequence to help transactivate the transcription of its target genes. One of the numerous target genes whose expression is regulated by NF-κB is, in fact, IκBα (Sun *et al.*, 1993; Scott *et al.*, 1993; Chiao *et al.*, 1994). The newly synthesized IκBα enters the nucleus and binds NF-κB and the resultant NF-κB- IκBα complex is expelled from the nucleus (Zabel *et al.*, 1993; Arenzana-Seisdedos *et al.*, 1995; Turpin *et al.*, 1999). This ability to export the NF-κB-IκBα complex from the nucleus is confered by the nuclear-export signal (NES) located in the N-terminus of the IκBα protein (Arenzana-Seisdedos *et al.*, 1997; Huang & Miyamoto, 2001). The IκB proteins retain NF-κB in the cytoplasm by masking nuclear-localization signals (NLSs) on NF-κB subunits (Ganchi *et al.*, 1992). In an NF-κB-IκBα complex, only one of the two NLSs in an NF-κB dimer is masked by IκBα, which allows the complex to shuttle constitutively between cytoplasm and nucleus of

quiescent cells (Malek *et al.*, 2001; Birbach *et al.*, 2002). By contrast, NF-κB-IκBβ complexes are sequestered in the cytoplasm because both NLSs on the NF-κB dimer are masked by IκBβ (Malek *et al.*, 2001).

Regulation of GRIA expression by NF-κB.

Putative NF-κB binding sites were identified by the MATCH program within the GRIA promoters. On careful investigation of the human GRIA promoters (Table 4.2), I found a distinct pattern among HSGRIA1, HSGRIA3 & HSGRIA4 in that they each have, at least, one NF-κB motif closely located (within 500 nt.) to and upstream of the TSS: in the HSGRIA1 promoter, this is found between –380 to –364; in HSGRIA3, it is between –151 to –136; and HSGRIA4 at –158 to –143 (and –29 to –14). The NF-κB motifs found downstream of the TSS in these 3 genes are not conserved between them: in HSGRIA3, the NF-kappaB motifs downstream of the TSS lie within the protein coding sequence. As for those motifs further upstream of the TSS, in the case of HSGRIA4, there are none found. For HSGRIA2, the putative NF-κB binding sites are all located very much upstream of the TSS (beyond 500nt. upstream of the TSS). Since GRIA2 endows the AMPA receptors with a different physiology by allowing AMPA receptors to be more Ca^{2+}-impermeable, its expression would most likely be slightly differently regulated from the other GRIAs. Therefore, it is reasonable to expect that the profile of the GRIA2 promoter is slightly different from those of the other GRIAs.

Prior studies documented increases in CNS tumor necrosis factor α (TNFα) within hours after the onset of ischemia (Liu *et al.*, 1994; Tseng & Chang, 1999; Gregersen *et al.*, 2000). Recent evidence suggests that TNFα sensitizes neurons to excitotoxic necrosis by inducing expression of GRIA1 via an acid sphingomyelinase (ASMase)- and NF-κB-dependent mechanisms (Yu *et al.*, 2002). NF-κB- or ASMase-deficient mice have been shown to be resistant to CNS injury after focal ischemia and coincidentally, the expression of GRIA1 is also reduced in the cerebral cortex and hippocampus of these mice (Schneider *et al.*, 1999; Yu *et al.*, 2000; Yu *et al.*, 2002). Using either an NF-κB p50 antisense oligonucleotide or a nonspecific ASMase inhibitor, desipramine, Yu and

colleagues (2002) found that they could inhibit TNFα-induced GRIA1 expression in cultured and differentiated NT2-N neurons. Furthermore, transient transfection of NT2-N neurons with NF-κB p50 induced expression of GRIA1 (Yu *et al.*, 2002). It was also found that the induction of GRIA1 in TNFα-treated NT2-N neurons increased their susceptibility to kainate necrosis (Yu *et al.*, 2002). Thus, we can clearly see there is strong evidence to suggest that TNFα increases the severity of ischemia-induced CNS necrosis, at least in part, by increasing the expression of GRIA1 through a NF-κB-dependent pathway.

IκBβ and its role in NF-κB activity.

It has been shown that TNFα induces a biphasic activation of NF-κB and it is believed that this biphasic activation is not cell type- or stimuli-specific (Thompson *et al.*, 1995; Ladner *et al.*, 2003; Schmidt *et al.*, 2003). In skeletal muscle, this biphasic activity consists of an initial phase of rapid but transient induction of NF-κB which peaked at 30 minutes post- TNFα treatment and returning to near basal levels by 1 hour (Ladner *et al.*, 2003). The second phase of NF-κB activity begins at around 4 hours later and persists for an additional 24-36 hours (Ladner *et al.*, 2003). In this case, it was found that the biphasic profile of NF-κB's activity was the result of nuclear translocation of NF-κB p65 which increased nuclear p65 levels (Ladner *et al.*, 2003). NF-κB is transcriptionally competent in both phases as demonstrated by a reporter-based assay and by analyzing the gene expression profiles of NF-κB responsive genes. The biphasic nature of TNFα-induced NF-κB activity might further be explained by differential rates of phosphorylation of IκBα and IκBβ by IKK: IKKβ was shown to be an efficient kinase for N-terminal serines of IκBα but phosphorylates the N-terminal serines of IκBβ far less efficiently, thereby explaining for the slower rate of degradation observed for IκBβ (Wu & Ghosh, 2003). Thus, the first phase of NF-κB activity might be attributed to the degradation of IκBα while the second phase might be the result of the slower degradation of IκBβ. In fact, the persistently increased NF-κB activity in the amnion during human labour is apparently due to the resistance of IκBβ-2 (one of two splicing isoforms of IκBβ) to degradation (Lee *et al.*, 2003). Unlike IκBα which is degraded and resynthesized rapidly and IκBβ-1

which is degraded more slowly as a result of IL-1b stimulation, IL-1b stimulation had little effect on IκBβ-2. Thus, despite an increased expression of inhibitory IκBα protein and the fact that no persistent IKK activity was detected, the only explanation for NF-κB's persistent activity might be IκBβ-2's resistance to degradation which might somehow function to protect NF-κB from inactivation in the amnion.

Suyang *et al.* (1996) had also reported observing a persistent, long-term induction of NF-κB activity in lipopolysaccharide (LPS)-treated B cells, much like that described for the second phase of the biphasic activation of NF-κB seen in TNFα-treated skeletal muscle. In B cells, the persistent activity was attributed to unphosphorylated, newly synthesized IκBβ binding to NF-κB and shielding it and preventing it from complexing with newly synthesized IκBα. Unphosphorylated IκBβ interacts with NF-κB differently to that of basally phosphorylated IκBβ because unphosphorylated IκBβ fails to mask the NLS and the DNA binding domain on NF-κB, it can complex with NF-κB in the cytoplasm and be imported into the nucleus or interact with NF-κB that are already bound to target promoters without displacing the NF-κB from the promoter, the result of which is a sustained NF-κB response in either case because, unlike IκBα, IκBβ does not have an NES at its N-terminus, which is essential for shuttling the NF-κB-IκBα complex out of the nucleus (Suyang *et al.*, 1996; Huang & Miyamoto, 2001). However, Suyang *et al.* (1996) found very low levels of unphosphorylated IκBβ in the nuclei of stimulated cells, suggesting that there might be an accompanying degradation of unphosporylated IκBβ in the nucleus.

Thus, the evidence above would suggest that IκBβ is responsible for the persistent activity of NF-κB. However, evidence to the contrary is also available. In human glial cells, endothelial cells, peritoneal macrophages and testis, a loss or reduction of IκBβ results in persistent NF-κB activity or the corollary that IκBβ expression inhibits NF-κB activity were shown (Johnson *et al.*, 1996; Velasco *et al.*, 1997; Bourke *et al.*, 2000; Budde *et al.*, 2002). IκBβ is highly expressed within the testis, more than any other tissue and this expression occurs in the virtual absence of IκBα expression (Budde *et al.*, 2002). IκBβ mRNA and protein expression is restricted to the haploid spermatid stages of

spermatogenesis and follows a wave of nuclear NF-κB expression within the earlier stages of spermatogenesis (Budde *et al.*, 2002). Given this, it would appear to be quite convincing that IκBβ serves to terminate NF-κB activity.

IκBβ is consitutively expressed within a number of cell types and tissues and unlike IκBα, its expression is not induced by NF-κB (Thompson *et al.*, 1995). Although a single NF-κB site is present within the IκBβ promoter that binds NF-κB, it can only modestly activate transcription of a reporter gene and was unable to strongly upregulate transcription in comparison to the NF-κB sites with the IκBα promoter (Budde *et al.*, 2002). However, it is possible that resynthesized IκBβ can be induced by other yet unknown pathways, for example, T cells treated with a synthetic peptide, DQ 65-79, is able to upregulate expression of IκBβ (Jiang *et al.*, 2002).

Considering the following facts presented above:

1. TNFα can sensitize neurons to excitotoxic necrosis by inducing expression of GRIA1 via an acid sphingomyelinase (ASMase)- and NF-κB-dependent mechanisms; and,

2. TNFα can induce a biphasic activation of NF-κB and that IκBβ plays a role in the second phase of this activation

it is likely that the persistent activation of NF-κB caused by IκBβ may result in a prolonged and increased expression of GRIA1. Although there is no convincing evidence to suggest that NF-κB is able to induce the transcription of IκBβ, we believe that common regulatory elements shared between IκBβ's and GRIA1's promoter might allow IκBβ's transcription to follow that of GRIA1 and feedback positively on the NF-κB-dependent GRIA1 transcription with the undesired consequence of increasing neuronal susceptibility to excitotoxic necrosis.

However, IκBβ's ability to terminate NF-κB activity under certain conditions (Johnson *et al.*, 1996; Velasco *et al.*, 1997; Bourke *et al.*, 2000; Budde *et al.*, 2002) would provide a

beneficial check to the NF-κB-dependent GRIA1 transcription under normal physiological conditions.

4.4.2 *Cellular trafficking and surface expression of AMPA receptors*

4.4.2.1 Delivery of AMPA rceptors to cell surface via exocytic pathways

It is obvious that the final destination of the AMPA receptor would be the cell surface where they would perform their function of neurotransmitter receptors. Thus, any gene/protein that aids the cellular trafficking of these receptors from the endoplasmic reticulum to the surface would play an important role in the surface expression of AMPA receptors and thus, would likely be co-regulated and/or co-expressed with the GRIAs.

The recruitment of AMPA receptors to the synapse occurs, for example, in activity-dependent synaptic plasticity at silent synapses and it is believed to be driven by calmodulin-dependent protein kinase II (CAMKII) (Isaac et al., 1995; Liao et al., 1995; Durand et al., 1996; Liao et al., 1999; Hayashi et al., 2000; Liao et al., 2001, Shi et al., 2001). LTP or increased CAMKII is shown to induce delivery of AMPA receptors into synapses of rat hippocampal neurons (Hayashi et al., 2000). Incidentally, CAMKII is both necessary and sufficient to generate calcium-evoked dendritic exocytosis (Maletic-Savatic et al., 1998). Brief activation of NMDA receptors in hippocampal slices can produce a long-lasting (>3 hours) increase in synaptic efficacy, that is, an NMDA-induced LTP (Broutman & Baudry, 2001). This NMDA-induced LTP also results in a rapid upregulation of GRIA1 and GRIA2/3 subunits in synaptic membranes

It is thought that the delivery of AMPA receptors to the synaptic membrane might be mediated by an exocytic pathway (Broutman & Baudry, 2001). A player in the vesicular and membrane fusion machinery, α-SNAP, while capable of enhancing the exocytic release of neurotransmitter in the squid giant synapse, has also been shown to enhance synaptic strength (DeBello et al., 1995; Lledo et al., 1998). In fact, on pathways in which LTP has been saturated, treatment with α-SNAP elicited only a slight increase in synaptic

strength (22 ± 12%) as compared with a control naïve pathway (58 ± 11%) (Lledo et al., 1998). Thus, it may be inferred that LTP occurs through the upregulation of AMPA receptors by a postsynaptic-regulated exocytic pathway employing a mechanism similar to that used for the release of neurotransmitters. Not surprisingly, botulinum toxin which disrupts the membrane fusion machinery by proteolytically cleaving the α-SNAP receptor, SNAP-25, can greatly reduce the magnitude of LTP (Lledo et al., 1998). Passafaro and colleagues (2001) further proposed that GRIA1 controls the exocytosis while GRIA2/3, the recycling and endocytosis of AMPA receptors.

In addition to their well-established postsynaptic action, AMPA receptors also mediate presynaptic effects (Nicoll et al., 2000). Presynaptic AMPA receptors have been shown to modulate synaptic transmission, by depressing the release of inhibitory GABA transmitters in the adult cerebellum (Satake et al., 2000). At steady state, a major pool of GRIA1 and GRIA2 subunits is associated with synaptic vesicle membranes (Schenk et al., 2003). Schenk and co-workers (2003) studying the delivery of AMPA receptors to the presynaptic membrane of axonal growth cones in hippocampal neurons, demonstrated that treatment with α-latrotoxin, which not only induces synaptic vesicle exocytosis but also prevents synaptic vesicle endocytosis in calcium-free cultured hippocampal neurons (Pennuto et al., 2002), gave significant staining of the growth cone with antibodies targeting the extracellular portion of the GRIA2 subunit. This implies that the massive fusion of synaptic vesicles was accompanied by the insertion of AMPA receptor subunits into the plasma membrane.

Below, I present evidence for 3 genes (VAMP4, Rab3B and FKBP8) that aid in the surface expression of AMPA receptors.

4.4.2.2 Vesicle-associated membrane protein 4 (VAMP4)

The secretory pathway compartments can be subdivided into 2 central membrane populations, the endoplasmic reticulum (ER)-Golgi system and the *trans*-Golgi network (TGN) system (Traub & Kornfeld, 1997; Gleeson *et al*., 2004). VAMP4 is broadly

expressed and localize to the Golgi-TGN (Advani *et al.,* 1998). VAMP4 protein was found in a complex with synaptophysin, physophilin, and VAMP1/3 in detergent extracts of rat brain membrane (Steegmaier *et al.*, 1999). Synaptophysin, the most abundant synaptic vesicle membrane protein known, is widely expressed throughout the nervous system and can also be found in chromaffin and neurosecretory granules (Jahn *et al.*, 1985; Wiedenmann & Franke, 1985; Buffa *et al.*, 1987; Schilling & Gratzl, 1988; Marqueze-Pouey *et al.*, 1991; Fykse *et al.*, 1993). Synaptophysin is known to form a complex with VAMP2 on the synaptic vesicle membrane during exocytosis (Calakos & Scheller, 1994; Edelmann *et al.*, 1995; Washbourne *et al.*, 1995; Galli *et al.*, 1996; Pennuto *et al.*, 2002). The use of fluorescent chimeras of synaptic vesicle proteins showed that synaptophysin is selectively confined to synaptic vesicles whereas other synaptic vesicle proteins, synaptotagmin, VAMP1 and VAMP2 were not exclusively localized to synaptic sites when overexpressed but were, instead, diffuse all over the surface of the axonal plasma membrane (Pennuto *et al.*, 2003). This means these synaptic vesicle proteins (synaptotagmin, VAMP1 and VAMP2), bear only the information that allows them to be sorted to the axon but not additional signals necessary for their recruitment to synaptic vesicles. Pennuto and co-workers (2003) showed that synaptophysin can selectively recruit VAMP2 to synaptic vesicles.

Synaptosomes purified from adult rat forebrain and stimulated with 0.1 nM α-latrotoxin in the absence of extracellular calcium showed an increase in cell surface GRIA2 and synaptophysin (Schenk et al., 2003). Synaptophysin, VAMP2, GRIA2/3 and GRIA1 could also be co-enriched in a vesicular fraction immunoisolated using magnetic beads coated with antibodies directed against synaptotagmin (Schenk et al., 2003). It is also interesting to note that Horikawa *et al.* (2002) identified γ-adaptin, a component of the AP-1 adaptor complex which recruits clathrin to membrane destined to form transport vesicles from the TGN, as a synaptophysin binding protein while, on the other hand, Eshhar and colleagues (1993) had previously found coated pits at the postsynaptic density that were immunoreactive for AMPA receptors in cultured hippocampal neurons (Seaman *et al.*, 1996; Robinson & Bonifacino, 2001). These experiments suggest that

synaptophysin may play a role in the formation of the clathrin-coated vesicles that transport the AMPA receptors to the synapse.

Ultrastructural studies have also shown VAMP4 on tubular and vesicular TGN membrane structures with a significant pool (31% of the total label) being found on clathrin-coated membranes (Steegmaier *et al.*, 1999). Bearing in mind the interaction of VAMP4 with synaptophysin (Steegmaier *et al.*, 1999) and synaptophysin's ability to selectively recruit VAMP2 to synaptic vesicles (Pennuto *et al.*, 2003), I believe that synaptophysin with the aid of VAMP4 might play a role in recruiting newly synthesized AMPA receptors to synaptic vesicles at the level of the TGN.

4.4.2.3 Rab3B (a member of the RAS oncogene family)

Rab3B belongs to the Rab3 family of small GTP-binding proteins which include Rab3A, Rab3C and Rab3D that function in regulated exocytosis (Touchot *et al.*,1987; Matsui *et al.*, 1988; Zahraoui *et al.*, 1989; Baldini *et al.*, 1992; Novick & Zerial, 1997; Lin & Scheller, 2000). The characterization of Rab3 protein family have been shown in different types of secretory cells that exhibit regulated exocytosis (Darchen *et al.*, 1990; Baldini *et al.*, 1992; Regazzi *et al.*, 1992; Weber *et al.*, 1996; Lin *et al.*, 1997; Tuvim *et al.*, 1999). Rab3B, for example, can potentiate the calcium-dependent secretion of noradrenaline when stably expressed in PC12 neuroendocrine cells (Weber *et al.*, 1996).

Rab3A, 3B and 3C mRNAs are expressed mostly in the brain (Moya *et al.*, 1992; Geppert *et al.*, 1994; Stettler *et al.*, 1995; Pavlos *et al.,* 2001; Schlüter *et al.,* 2002). Rab3D, also implicated in exocytosis, is expressed primarily outside the brain in exocrine glands and in mast cells where it is enriched on secretory vesicles, although very low levels of the protein were detected in the brain (Ohnishi *et al.*, 1996; Valentijn *et al.*, 1996a, b; Tuvim *et al.*, 1999; Schlüter *et al.,* 2002). In adrenal chromaffin cells, anti-Rab3B antibody staining was found on the plasma membrane (Lin *et al.*, 1997). However, in the brain, evidence indicates that Rab3B (along with Rab3A and 3C) are highly concentrated on synaptic vesicles (Fischer von Mollard *et al.*, 1990; Schlüter *et al.,* 2002). In fact, using a

stringent purification method for synaptic vesicles, it was shown that Rab3B could be co-purified with Rab3A and synaptophysin, thus indicating that Rab3B is a synaptic vesicle protein (Schlüter *et al.,* 2002).

As mentioned above, recruitment of AMPA receptors to the synapse is believed to be driven by calmodulin-dependent protein kinase II (CaMKII) (Isaac *et al.*, 1995; Liao *et al.*, 1995; Durand *et al.*, 1996; Liao *et al.*, 1999; Hayashi *et al.*, 2000; Liao *et al.*, 2001, Shi *et al.*, 2001). Rab3B binds and interacts with Ca^{2+}-calmodulin (CaM) in a calcium-dependent manner with a reduction in binding seen in the presence of EDTA (Sidhu & Bhullar, 2001). Rab3B's involvement in Ca^{2+}-dependent exocytosis has been explicitly shown in rat anterior pituitary cells (Lledo *et al.*, 1993). By use of Rab3B antisense, it was found that Ca^{2+}-dependent exocytosis was inhibited as measured by changes in the membrane capcitance. Rab3A antisense, on the other hand, had no effect on Ca^{2+}-dependent exocytosis (Lledo *et al.*, 1993).

The Rab3 proteins seem to function similarly in exocytosis and this coupled with the finding that deletion of Rab3A in knock-out mice leads to a relatively mild phenotype that includes altered short-term synaptic plasticity and the absence of a presynaptic form of LTP have led some to think that this may be a consequence of redundancy among Rab3 isoforms (Geppert *et al.,* 1997; Castillo *et al.,* 1997). However, Schlüter and co-workers (2002) found although Rab3A, Rab3B, and Rab3C are co-localized on synaptic vesicles, they exhibit a differential distribution among the brain regions. Rab3A is uniformly present in all brain areas while Rab3C is found in most brain areas at variable levels but Rab3B is only found in a subset of brain areas. Rab3B mRNA was found abundantly in the olfactory bulb (in the mitral cell layer) and within the pituitary (particularly within the anterior and intermediate lobes) and this is confirmed with immunoblotting analyses (Stettler *et al.*, 1995; Schlüter *et al.*, 2002). Rab3B expression was also detected in the hippocampus being particularly prominent in the dentate gyrus and CA1 region with the CA3-CA4 region showing a lower level of expression (Stettler *et al.*, 1995; Schlüter *et al.*, 2002). Among other distinct regions of the brain where Rab3B could be found at moderate to high levels were the thalamus, the piriform cortex,

the hypothalamus (especially within the supraoptic and pariventricular nuclei) and the cerebral cortex. Weak Rab3B expression were obeserved in the caudate putamen and the Purkinje cells of the cerebellum (Stettler *et al.*, 1995). Careful look at the distribution of Rab3B shows that there is considerable overlap in expression with the AMPA receptor subunits in both adult and developing brains (Hollmann *et al.*, 1989; Boulter *et al.*, 1990; Keinänen *et al.*, 1990; Pellegrini-Giampietro *et al.*, 1991; Rogers *et al.*, 1991; Martin *et al.*, 1993; Hollmann & Heinemann, 1994). Most striking is the high levels of both Rab3B and AMPA receptors in the hippocampus, where the expression of the GRIA subunits are well-documented in the dentate gyrus, CA1 and CA3 regions. Also in the pituitary, electrophysiological studies have identified the presence functional AMPA-kainate glutamate receptors in the same lobes (the anterior and intermediate lobes) where Rab3B is expressed (Stettler *et al.*, 1995; Poisbeau *et al.*, 1996; Villalobos *et al.*, 1996). Immunocytochemistry studies using specific antibodies to the AMPA receptor subunits also confirmed the presence of GRIA1 and GRIA2/3-positive cells in the anterior and intermediate lobes of the pituitary (Kiyama *et al.*, 1993).

The localization of Rab3B on synaptic vesicles and its demonstrated functional role in exocytosis, coupled with its high expression in the brain, particularly the matching expression profiles with AMPA glutamate receptor subunits, provide a strong indication that Rab3B might be enlisted to help in the surface expression of AMPA glutamate receptors.

4.4.2.4 FK506 binding protein 8, 38kDa (FKBP8)

FK506 binding protein 8 (FKBP8) is a member of the immunophilin protein family (Lam *et al.*, 1995; Pedersen *et al.*, 1999). Little work has been carried out with FKBP8 but FKBP8 is 33% identical to FKBP12 in the N-terminus (between aa 44 and 142) with a consensus leucine zipper in a predicted α-helical region (Lam *et al.*, 1995). FKBP8 also has a 3-unit tetratricopeptide repeat (TPR) domain and in its extreme C-terminus, a putative CaM-binding site (Lam *et al.*, 1995). Immunophilins are a family of receptor proteins for the immunosuppresant drugs cyclosporin A (CsA), FK506 and rapamycin,

which are used during organ transplantation (Schreiber, 1991). Research on immunophilins had, understandably, been focused on its role in cells of the immune system, especially lymphocytes (Schreiber, 1991; Sigal & Dumont, 1992). However, levels of FKBP12 in the brain were found to be up to 50 times greater than those in tissues of the immune system, suggesting a neural role for the immunophilins (Steiner *et al.*, 1992). Moreover, the distribution of FKBP12 and cyclophilin in the brain is almost exclusively neuronal with marked regional variations that, coincidentally, closely resemble the distribution of calcineurin (Dawson *et al.*, 1994).

As explained above, recruitment of AMPA receptors to the synapse involves exocytic pathways similar to those involved in neurotransmitter release. Incidentally, FK506 blocks NMDA-induced release of glutamate from brain synaptosomes (Steiner *et al.*, 1996). In addition, FK506 also prevents spontaneous and K^+-depolarization-induced neurotransmitter release from PC12 cells (Steiner *et al.*, 1996). It has been shown that the FK506-FKBP12 complex interacts with calcineurin, a CaM-activated protein phosphatase, to inhibit its phosphatase activity (Liu *et al.*, 1991). As a result, it is believed that the drug-immunophilin complex indirectly maintains neuronal nitric oxide synthase (nNOS) in its phosphorylated state and thus, reducing its catalytic activity, leading to a fall in NO levels (Dawson *et al.*,1993). NO is required for neurotransmitter release in PC12 cells and NMDA-stimulated synaptosomes, since a similar block in neurotransmitter release can be observed with NOS inhibitors in these systems (Hirsch *et al.*, 1993). Glutamate release has also been shown to be induced by NO in other systems (Bal-Price & Brown, 2001; Matsuo *et al.,* 2001). Thus, we see here that immunophilins can play a regulatory role in neurotransmitter release, with calcineurin indirectly promoting neurotransmitter release while the drug-immunophilin complex inhibiting this release through an inhibition of calcineurin.

On the other hand, the pathway responsible for the AMPA receptor upregulation might be different from those for NMDA-stimulated glutamate release, even if both occur through exocytic mechanisms. NMDA-induced LTP saw a rapid upregulation of GRIA1 and GRIA2/3 subunits in synaptic membranes (Broutman & Baudry, 2001). KN-62, a

CaMKII inhibitor, completely inhibits this NMDA-induced upregulation of GRIA1 and GRIA2/3 subunits and consequently, NMDA-induced LTP. Calcineurin, being a phosphatase, might act similarly to KN-62 since it might nullify the actions of CaMKII. Both calcineurin and CaMKII are activated by calmodulin, but it is possible that immunophilins (with an endogenous ligand) might act to inhibit calcineurin, while CaMKII helps upregulate the expression of AMPA receptors.

Heat-shock protein 90 (hsp90) was recently shown to be a critical component required for the constitutive trafficking of AMPA receptors into synapses during their continuous cycling between synaptic and non-synaptic sites (Gerges *et al.*, 2004). Radicicol, a hsp90-specific inhibitor, prevents the constitutive delivery of AMPA receptors, as assayed with recombinant GRIA2, and decreases AMPA receptor-mediated responses. Not surprisingly then, hsp90 is expressed constitutively in brain from early development into adulthood (D'Souza & Brown, 1998). It has been shown that hsp90 is required for the subcellular targeting of a variety of receptor proteins, including the glucocorticoid receptor (Owens-Grillo *et al.*, 1996; Czar *et al.*, 1997; Silverstein *et al.*,1999; Galigniana *et al.*, 2001), the dioxin receptor (Kazlauskas *et al.*, 2000; Kazlauskas *et al.*, 2001), the receptor tyrosine kinse ErbB2 (Xu *et al.*, 2002) and the epidermal growth factor receptor (Supino-Rosin *et al.*, 2000). The trafficking functions of hsp90 involve specific interactions between C-terminal sequences of hsp90 and TPR domains in several effector molecules (Chen *et al.*, 1996; Young *et al.*, 1998; Russell *et al.*, 1999; Scheufler *et al.*, 2000; Ward *et al.*, 2002). Overexpression of the TPR domain of protein phosphatase 5, which binds specifically to hsp90, significantly decreased AMPA receptor-mediated transmission without altering NMDA-mediated responses in CA1 hippocampal neurons (Gerges *et al.*, 2004). In contrast, expression of the TPR domain did not alter LTP induction, suggesting that hsp90 is not involved in the activity-dependent delivery of AMPA receptors. Thus, hsp90's function in trafficking AMPA receptors to the synapse may be mediated by a TPR-domain containing protein. FKBP8 has a 3-unit TPR domain (Lam *et al.*, 1995).

In a hsp90-glucocorticoid receptor heterocomplex, FKBP52 (otherwise known as FKBP4 or FKBP59) was found to bind directly to hsp90 via its 3 TPR domains (Renoir *et al.*, 1990; Radanyi *et al.*, 1994; Czar *et al.*, 1994; Bermingham *et al.*, 1998). It is suggested that FKBP52 helps target receptor movement to the nucleus in this case (Pratt *et al.*, 1993). This binding site can be competed by another immunophilin, cyclophilin 40 (CyP40) and the two immunophilins exist in independent cytosolic heterocomplexes with hsp90 and with the untransformed glucocorticoid receptor (Owens-Grillo *et al.*, 1995; Owens-Grillo *et al.*, 1996; Ratajczak & Carrello, 1996). Through a series of binding studies with other TPR-containing proteins, Owens-Grillo and colleagues (1996) postulated that hsp90 has a universal TPR domain-binding region that permits it to bind to multiple proteins. FKBP8, like FKBP52 and CyP40, has 3 TPR domains and therefore, is very likely to also bind hsp90 (Lam *et al.*, 1995). It is very likely that FKBP8 forms a receptor heterocomplex with hsp90 and AMPA receptors and play a role in the constitutive delivery of AMPA receptors to the synapse.

Coincidentally, hsp90 was shown to be necessary for efficient neurotransmitter release at the presynaptic terminal (Sakisaka *et al.*, 2002; Gerges *et al.*, 2004). This lends additional support to the notion that the machinery responsible for neurotransmitter release could also be responsible or, at least, be similar to the machinery for delivery of AMPA receptors to the synapse.

4.4.3 Morphogenesis

4.4.3.1 Involvement of AMPA receptors in development and dendritogenesis

The idea that synaptic activity may influence the wiring of the brain, in particular, dendritogenesis, is confirmed in various developmental situations (Katz & Constantine-Paton, 1988; Rajan & Cline, 1998; Sin *et al.*, 2002). The role of glutamate receptors in the development of the nervous system is well-established (Molnar *et al.*, 2002). It has been proposed that neural activity-induced changes in neural circuitry, such as in activity-dependent development or LTP, may involve so-called “silent synapses” (Liao *et al.*,

1995; Isaac *et al.*, 1995; Durand *et al.*, 1996; Liao *et al.*, 1999). That is, dendrites bearing postsynaptic NMDA glutamate receptors (GRINs) make a significant number of synaptic contacts with the axonal presynaptic membrane but they remain postsynaptically "silent" at resting potential because of the voltage-dependent blockade of GRINs by magnesium until such time when AMPA receptors are acquired, for example, after a LTP-inducing protocol that last only minutes; branches that acquire GRIAs are stabilized and those that do not are retracted (Mayer *et al.*, 1984; Nowak *et al.*, 1984; Isaac *et al.*, 1995; Liao *et al.*, 1995; Liao *et al.*, 2001). It is interesting to note that Hohnke *et al.* (2000) found that although a small number of silent retinogeniculate synapses are present, there is no overall change in GRIA/ GRIN contribution when the retinogeniculate axons from ON-center and OFF-center retinal ganglion cells segregate to form ON/OFF sublaminae in the lateral geniculate nucleus. Thus, they argue against the idea that the conversion of silent to functional synapses could play a role in the development and refinement of inputs. It would be naive to believe that one mechanism could serve all cases of activity-dependent development. Moreover, what is overlooked and the conclusion that can be drawn from all these studies is the fact that although axon guidance and development may not be dependent on GRIA, but perhaps GRIA-dependent mechanisms may play a role in dendritogenesis.

It was recently shown that expression of GRIA1flip alone in architecturally mature dendrites is sufficient to initiate a remodeling of the dendritic arbor (Inglis *et al.*, 2002). The repertoire of glutamate receptors expressed by developing motor neurons differs significantly from the glutamate receptor phenotype of mature motor neurons (Kalb *et al.*, 1992; Stegenga & Kalb, 2001). Neonatal motor neurons express very high levels of the GRIA1flip subunit but not adult motor neurons (Jakowec *et al.*, 1995a, b). Inglis and colleagues (2002) were able to show that GRIA1flip might be involved in the modeling of dendritic architecture of motor neurons during development and possibly, this mechanism, with the appropriate stimulus, can be engaged in mature motor neurons to modify the existing dendritic architecture. Interestingly, Zamanillo *et al.* (1999) discovered that in GRIA1$^{-/-}$ adult mice, associative LTP was absent in hippocampal CA3 to CA1 synapses.

With evidence pointing to AMPA receptors' involvement in the formation and change of neural circuitry, it is natural to suspect that certain genes / proteins involved in cellular morphogenesis would also be co-expressed along with the GRIA genes to aid in this functional role of the AMPA receptors. Below we look at two genes, 3-OST-3_A and CLSTN3, and provide supporting evidence for our supposition that these genes are co-regulated / co-expressed along with the GRIAs.

4.4.3.2 Heparan Sulfate D-glucosaminyl 3-*O*-sulfotransferase 3A (3-OST-3_A)

Heparan sulfate proteoglycans (HSPGs) are implicated in various cellular scenarios including cell proliferation, adhesion, migration, and morphogenesis during development (Syrokou *et al.*, 1999; Inatani *et al.*, 2003; Kaneider *et al.*, 2004; Beauvais & Rapraeger, 2004) and in various physiological conditions such as inflammation, blood coagulation and tumour malignancy (Tanaka *et al.*, 1993; Shworak *et al.*, 1996; Harada *et al.*, 2003). The heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase multigene family of enzymes is a group of enzymes that generates these highly specific sulfated oligosaccharide sequences in HSPGs (Liu *et al.*, 1996; Liu *et al.*, 1999a, b; Shworak *et al.*, 1999). Heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase (3-OST-1), for example, is a key enzyme converting nonanticoagulant heparan sulfate (HS) to anticoagulant HS (Liu *et al.*, 1996). Each enzyme in this family has novel but distinct substrate specificities and it was demonstrated that, unlike 3-OST-1, 3-OST-3_A is very much (300 times) less effective in converting HS to its active anticoagulant form, suggesting that it does not make anticoagulant HS (Liu *et al.*, 1999a, b).

Heparan sulphate (HS) is expressed abundantly and in a regulated manner during development in the mammalian CNS, suggesting a functional role in brain development (Yamaguchi, 2001). Newborn mice in which HS synthesis in the brain is disrupted does not survive past the first day of birth (Inatani *et al.*, 2003) Despite a normal gross appearance of the embryo, mice with disrupted HS synthesis in the brain showed specific developmental defects in their CNS: most notably a patterning (malformation) defect in

the midbrain-hindbrain region, characterized by the absence of a discernible inferior colliculus and cerebellum, a phenotype similar to that caused by a hypomorphic *Fgf8* allele and a natural *Wnt1* allele called swaying (Thomas *et al.*, 1991b; Meyers *et al.*, 1998; Inatani *et al.*, 2003). As it turns out, the lack of HS was found to disturb the expression of downstream genes. For example, the distribution of FGF8 was expanded, unlike the concentrated FGF8 band typically seen in the midbrain-hindbrain boundary of normal wild-type embryos. In turn, other downstream genes such as *Wnt1*, *Engrailed 1* (*En1*) and *Engrailed 2* (*En2*) also showed a abnormal expression profile in the brain. Interestingly, immunostaining for calbindin, a marker for Purkinje cells, revealed a disorganized collection of calbindin-positive cells, suggesting that the specification of cerebellum occurred but subsequent developmental steps leading to the formation of a fully organized inferior colliculus and cerebellum were halted (Inatani *et al.*, 2003).

Liitle research has been done with 3-OST-3_A but in our study, we found that the promoter of 3-OST-3_A share certain features with those of GRIAs, which lead us to believe that the expression of 3-OST-3_A and GRIAs is co-regulated. We propose that when 3-OST-3_A is expressed in GRIA-containing cells, it might act to catalyze HSPGs to an active form that is crucial for brain morphogenesis, for example, in the modeling of dendritic architecture.

Paradoxically, Chisamore *et al.* (1996) found that while proteoglycan release is generally increased in glutamate-treated fetal hippocampal astrocytes, the release of HSPG to the extracellular enviroment of the astrocyte was reduced. Studies with cultured and mature hippocampal astrocytes have revealed the presence of AMPA/kainate and metabotropic receptors, while also indicating the absence of NMDA receptors (Condorelli *et al.*, 1993; Shelton & McCarthy, 1999). AMPA receptors and HSPGs have a role in CNS development and if 3-OST-3_A converts HSPGs to an active form necessary for brain morphogenesis, it would not make much sense for 3-OST-3_A to be co-regulated/expressed with GRIAs since the end-effect of glutamate activation is inhibiting the release of HSPG which is counterproductive to the actions of 3-OST-3_A. However, it is postulated that while HSPG is necessary for brain morphogenesis, GRIAs' role here

might be to provide a feedback loop during this process. That is, for example, in activity-dependent development, a GRIA-bearing dendrite may release active HSPG to facilitate its development but, on contact with the appropriate presynaptic terminal, it would receive glutamatergic stimulation to inhibit the release of active HSPG.

4.4.3.3 Calsyntenin 3 (CLSTN3)

Calsyntenin 3 belongs to a family of novel postsynaptic membrane proteins that were discovered only recently (Vogt *et al.*, 2001; Hintsch *et al.*, 2002). The discovery of this family of proteins was made on the back of attempts to identify target proteins of extracellular proteases. Synaptic plasticity is accompanied by structural changes in and around the synapse and these structural reorganizations were confirmed to be due to the action of extracellular proteases (Engert & Bonhoeffer, 1999; Toni *et al.*, 1999; Shiosaka &Yoshida, 2000; Yuste & Bonhoeffer, 2001; Oka *et al.*, 2002). Extracellular proteases such as tissue plasminogen activator (tPA) and neuropsin were shown to play a role in synaptic plasticity (Huang *et al.*, 1996; Frey *et al.*, 1996; Baranes *et al.*, 1998; Hirata *et al.*, 2001; Mataga *et al.*, 2002; Matsumoto-Miyai *et al.*, 2003). For instance, targeting of the tPA gene in mice resulted in selective interference of the late-phase LTP in the hippocampus and was characterized by stronger GABAergic transmission in the hippocampal CA1 region (Frey *et al.*, 1996; Huang *et al.*, 1996). Little work has been carried out on calsyntenins and thus, exact function of this family of proteins is unknown.

However, it was found that calsyntenins are predominantly expressed in the neurons (Vogt *et al.*, 2001; Hintsch *et al.*, 2002). Immunoelectron microscopy reveals that all three calsyntenins are located in the postsynaptic membrane of asymmetrical (excitatory) synapses (Hintsch *et al.*, 2002). There is ample evidence showing that excitatory AMPA glutamate receptors are also expressed on postsynaptic dendritic spines of asymmetrical synapses (Baude *et al.*, 1995; Bernard *et al.*, 1997; Clarke & Bolam, 1998; Bernard & Bolam, 1998; Kessler & Baude, 1999). In addition, Hintsch and co-workers (2002) also demonstrated that several clearly discernible populations of glutamatergic neurons, such

as those in neocortical layer 5 and the hippocampal CA1-CA3 regions, expressing high levels of CLSTN3.

The fact that CLSTN3 is a target of extracellular proteases which play a central role in synaptic plasticity coupled with the fact that its expression pattern indicates that it is expressed in cells which also express AMPA receptors, supports the belief that the similar promoter profile of these two genes is of no coincidence and that calsyntenin 3 and AMPA receptors are co-expressed /co-regulated in cases of where neurons undergo synaptic plasticity, such as in LTP.

Table 4.1

List of genes that share the unique promoter profile found in GRIA promoters (Locus ID refers to the designated NCBI LocusLink identification number for the gene)

Genes coding for proteins involved in receptor trafficking and surface expression				
	Locus ID	Name	Symbol(s)	References
1.	8674	Vesicle-associated membrane protein 4	VAMP4, VAMP24	Advani *et al.,* 1998; Steegmaier *et al.*, 1999
2.	23770	FK506 binding protein 8, 38kDa	FKBP8, FKBP38	Lam *et al.*, 1995; Pedersen *et al.*, 1999
3.	5865	RAB3B: member of RAS oncogene family	RAB3B,	Zahraoui *et al.*, 1989 Schlüter *et al.,* 2002
Genes coding for proteins involved in cellular morphology and structure				
	Locus ID	Name	Symbol(s)	References
1.	9955	heparan sulfate (glucosamine) 3-O-sulfotransferase 3A1	HS3ST3A1, 30ST3A1, 3-OST-3_A	Liu *et al.*, 1999a; Liu *et al.*, 1999b
2	79690	galactose-3-O-sulfotransferase 4	GAL3ST4 FLJ12116	Chandrasekaran *et al.*, 2004
3.	9746	Calsyntenin 3	CLSTN3, CSTN3, alcbeta, KIAA0726	Vogt *et al.*, 2001; Hintsch *et al.*, 2002
4.	10097	Actin-related protein 2 homolog	ARP2, ACTR2	Li *et al.*, 2004; Lynch *et al.*, 2003; Yarar *et al.*, 2002.
Genes coding for transcription factors (TF) and proteins regulating TF function				
	Locus ID	Name	Symbol(s)	References
1.	8651	Suppressor of cytokine signaling 1	SOCS1, SSI-1, CIS1, CISH1, JAB, TIP3	Starr *et al.*, 1997; Krebs & Hilton, 2000
2.	4793	Nuclear factor of kappa light polypeptide gene	NFKBIB,	Thompson *et*

		enhancer in B-cells inhibitor, beta	IκBβ, TRIP9	*al.*, 1995; Li & Verma, 2002
3.	9457	Activator of cAMP-responsive element modulator (CREM) in testis	ACT, FLJ33049, dJ393D12.2	Fimia *et al.*, 2001; Fimia *et al.*, 1999
4.	26959	HMG-box transcription factor 1	HBP1	Swanson *et al.*, 2004; Sampson *et al.*, 2001; Tevosian *et al.*, 1997
5.	83595	SRY (sex determining region Y)-box 7	SOX7, MGC10895	Takash *et al.*, 2001
6.	7707	Zinc finger protein 148	ZNF148, BERF-1, BFCOL1, ZBP-89, ZFP148, pHZ-52, HT-BETA	Wu *et al.*, 2004; Law *et al.*,1998; Wang *et al.*, 1993
Genes coding for proteins involved in protein biosynthesis				
	Locus ID	Name	Symbol(s)	References
1.	6169	Ribosomal protein L38	RPL38	Yoshihama *et al.*, 2002; Espinosa *et al.*, 1997
2.	29074	Mitochondrial ribosomal protein L18	MRPL18, HSPC071	Zhang & Gerstein, 2003; Koc *et al.*, 2001
Genes coding for proteases				
	Locus ID	Name	Symbol(s)	References
1.	5653	Kallikrein 6 (also known as Neurosin or Zyme)	KLK6, Bssp, Klk7, SP59, PRSS9, PRSS18, MGC9355	Scarisbrick *et al.*, 2002; Yousef *et al.*, 1999; Anisowicz *et al.*, 1996
2.	56999	A Disintegrin-like and metalloprotease with thrombospondin type 1 motif, 9	ADAMTS9, KIAA1312	Clark *et al.*, 2000.
3.	11131	Calpain 11	CAPN11	Dear *et al.*, 1999

Genes coding for proteins involved in lipid metabolism				
	Locus ID	Name	Symbol(s)	References
1.	341	Apolipoprotein C-I	APOC1	Mehta *et al.*, 2003; Ki *et al.*, 2002
2.	5279	Phosphatidylinositol glycan, class C	PIG-C, GPI2, MGC2049	Eisenhaber *et al.*, 2003; Watanabe *et al.*, 1998; Inoue *et al.*, 1996
Genes coding for secreted protein(s)				
	Locus ID	Name	Symbol(s)	References
1.	27120	Soggy (Dickkopf-like 1)	DKKL1 SGY, SGY1	Krupnik *et al.*, 1999; Clark *et al.*, 2003.
Genes coding for mitochondrial protein(s)				
	Locus ID	Name	Symbol(s)	References
1.	10063	Human COX17 homolog (cytochrome c oxidase assembly protein [yeast])	COX17	Punter *et al.*, 2000; Horvath *et al.*, 2000; Amaravadi *et al.*, 1997.
Genes coding for kinase(s)				
	Locus ID	Name	Symbol(s)	References
1.	4921	Discoidin domain receptor family, member 2	DDR2, TKT, NTRKR3, TYRO10	Ferri *et al.*, 2004; Vogel *et al.*, 1997; Karn *et al.*, 1993.
Genes coding for proteins related to retinoic acid and its metabolism				
	Locus ID	Name	Symbol(s)	References
1.	5919	Retinoic acid receptor responder 2	RARRES2, TIG2, HP10433	Nagpal *et al.*, 1997.
2.	64577	aldehyde dehydrogenase 8 family, member A1	ALDH8A1, ALDH12, DJ352A20.2	Lin & Napoli, 2000.
Genes coding for hormones				
	Locus ID	Name	Symbol(s)	References
1.	6013	Relaxin 1	RLN1, RLXH1, bA12D24.3.1, bA12D24.3.2	Garibay-Tupas *et al.*, 1999; Hudson *et al.*, 1983.
Genes coding for proteins involved in cell growth				

	Locus ID	Name	Symbol(s)	References
1.	11170	TU3A protein	TU3A, DRR1	Wang *et al.*, 2000; Yamato *et al.*, 1999
Genes coding for proteins involved in angiogenesis				
	Locus ID	Name	Symbol(s)	References
1.	576	Brain-specific angiogenesis inhibitor 2	BAI2	Kee *et al.*, 2002; Shiratsuchi *et al.*, 1997.
Genes coding for proteins involved in immunity and inflammation				
	Locus ID	Name	Symbol(s)	References
1.	2323	fms-related tyrosine kinase 3 ligand	FLT3LG,	Agrawal *et al.*, 2001; Hannum *et al.*, 1994; Lyman *et al.*, 1994.
2.	3603	Interleukin 16	IL-16, LCF, prIL-16, HsT19289	Kurschner & Yuzaki, 1999; Baier *et al.*, 1997; Cruikshank *et al.*, 1994.
3.	911	CD1C antigen, c polypeptide	CD1C, CD1	Zheng *et al.*, 2002; Martin *et al.*, 1987.
Genes coding for proteins with unknown function				
	Locus ID	Name	Symbol(s)	References
1.	80206	hypothetical protein FLJ22297	FLJ22297, KIAA1695, FHOD3, FHOS2, FLJ22717	Katoh & Katoh, 2004; Nagase *et al.*, 2000.
2.	80227	hypothetical protein FLJ11848	FLJ11848	Strausberg *et al.*, 2002.
3.	84217	zinc finger, MYND domain containing 12	ZMYND12 DKFZp434N2435	Strausberg *et al.*, 2002.
4.	22845	transmembrane protein 15	TMEM15 KIAA1094	Clark *et al.*, 2003; Strausberg *et al.*, 2002.
5.	23276	kelch-like 18 (Drosophila)	KLHL18 FLJ13703 KIAA0795	Strausberg *et al.*, 2002; Nagase *et al.*, 1998.
6.	25852	armadillo repeat containing 8	ARMC8	Strausberg *et*

			HSPC056 MGC4880 MGC10058 DKFZP434A043	*al.*, 2002.
7.	29082	chromosome 14 open reading frame 123	C14orf123 Shax2 CHMP4A CHMP4B HSPC134	Strausberg *et al.*, 2002; Zhang *et al.*, 2000.
8.	54093	chromosome 21 open reading frame 18	C21orf18	Strausberg *et al.*, 2002; Reymond *et al.*, 2001; Hattori *et al.*, 2000.
9.	54897	hypothetical protein FLJ20321	FLJ20321	Strausberg *et al.*, 2002.
10.	55244	hypothetical protein FLJ10847	FLJ10847	Bi *et al.*, 2002; Strausberg *et al.*, 2002.
11.	55727	BTB (POZ) domain containing 7	BTBD7 FUP1 MGC48310 FLJ10648	Pan *et al.*, 2001; Strausberg *et al.*, 2002.
12.	55859	brain expressed, X-linked 1	BEX1 HBEX2 HGR74-h	Yang *et al.*, 2002.
13.	64225	ADP-ribosylation-like factor 6 interacting protein 2	ARL6IP2 FLJ23293	Strausberg *et al.*, 2002.
14.	79740	hypothetical protein FLJ23049	FLJ23049	Strausberg *et al.*, 2002.
15.	79980	chromosome 20 open reading frame 172	C20orf172 FLJ13346, dJ469A13.2	Strausberg *et al.*, 2002.
16.	83892	potassium channel tetramerisation domain containing 10	KCTD10 MSTP028	Strausberg *et al.*, 2002.

Table 4.2

NF-κB motifs found by MATCH within the promoters of human GRIA genes

NF-κB	Positions (wrt TSS)		Strand
HS GRIA1	1	-630 to -615	-ve
	2	-503 to -488	+ve
	3	-380 to -365	-ve
	4	-379 to -364	-ve
	5	80 to 95	+ve
	6	81 to 96	+ve
	7	249 to 264	-ve
	8	388 to 403	-ve
HS GRIA2	1	-1250 to -1235	-ve
	2	-1249 to -1234	-ve
	3	-1163 to -1148	+ve
	4	-1030 to -1015	+ve
	5	-803 to -788	+ve
	6	-731 to -716	-ve
	7	-656 to -641	-ve
	8	813 to 828	+ve
HS GRIA3	1	-1129 to -1114	+ve
	2	-151 to -136	+ve
	3	313 to 328	-ve
	4	402 to 417	-ve
	5	679 to 694	-ve
HS GRIA4	1	-158 to -143	-ve
	2	-29 to -14	-ve
	3	219 to 234	-ve
	4	291 to 306	+ve
	5	966 to 981	-ve

CHAPTER 5: CONCLUDING REMARKS

The primary goal of the thesis has been the identification key transcriptional elements that control the expression of AMPA receptor genes. I wanted to ask the question "What is unique to the promoters of AMPA receptor genes and not commonly found in the promoters of other human genes?". To this end, it was first important to develop a software that could automate the collection of human promoter sequences with a high degree of accuracy and therefore, resulted in the development of the 5'-end Information Extraction (FIE) system. With FIE version 2, it was possible to collect some 10,000-odd human promoter sequences covering 1500 upstream to 1000 downstream of the identified TSS.

Following this, a novel bioinformatics approach recently developed by Bajic *et al.* (2004) was applied to study the AMPA receptor promoters. The human AMPA receptor is made by a combination of one or more of 4 subunits, GRIA1-GRIA4. For this study, available promoter sequences of the mouse AMPA receptor subunits and rat GRIA1 subunit were also included. The rationale for this is because it was believed that if a transcription factor binding site (TFBS) or composite element was truly essential to the regulation of AMPA receptor gene regulation, then it should naturally be conserved across the species. With this approach, a combination of the top 3-ranked "singles", "pairs" and "triplets could describe well the AMPA receptor promoters was discovered. This combination of TFBSs and composite elements was not entirely unique to the AMPA receptor promoters since 47 other human gene promoters (out of a pool of 10,741 human promoters) were found to contain this same combination. However, the uniqueness of this combination of TFBSs and composite elements is still statistically significant considering that it is represented in less than 0.5% of a comprehensive group of gene promoters. Furthermore, I hypothesize that the expression of the 47 genes are coregulated with AMPA receptor genes and this would explain their shared promoter profile. It is often said that there are two sets of transcriptional elements within a gene's promoter: one which is gene-specific and another which is tissue-specific. However, in this case, it would seem that the combination of TFBSs and composite elements that has been identified within the AMPA receptor gene

promoters is function-specific, that is to say, this particular combination of TFBSs and composite elements control the expression of genes that aid in the expression of AMPA receptors and its functioning. On the other hand, the phylogenetic footprinting study on the human, mouse and rat GRIA1 gene yields a gene-specific set of transcriptional elements.

On should recognize that all bioinformatics tools generate false positive results and are not 100% foolproof, however, the approach used is statistically sound and the above hypothesis is supported by strong experimental evidence. The method of feature identification used greatly reduces false positive errors normally generated by the conventional TFBS prediction programs such as MATCH and provides a higher level of statistical significance because it is not identifying individual TFBSs or composite elements on a single promoter sequence but the combination of regulatory elements that is uniquely characteristic to several promoter sequences in a gene family. Thus, I am confident of the results generated by this approach.

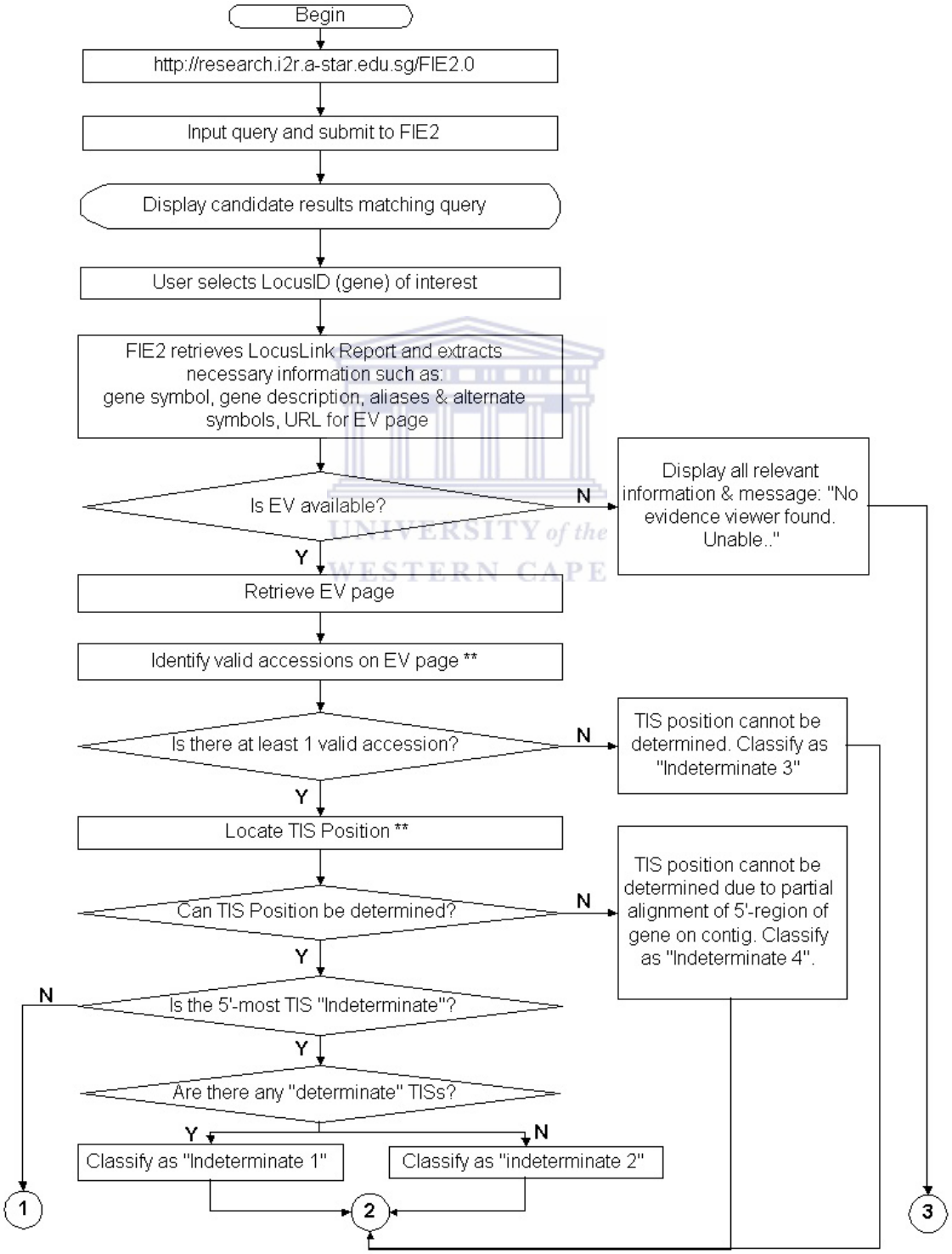
Here, evidence is also given to show how seven genes which had no prior relationship to AMPA receptors are, in fact, key players in AMPA receptor physiology; thus, explaining why their expression is coregulated with the GRIA genes. The role of these seven genes in AMPA receptor physiology is not always immediately evident and they seem to be involved in everything from transcriptional regulation, such as SOCS1, to synaptic plasticity and morphogenesis, such as 3-OST-3_A. The elucidation of the role of these genes in AMPA receptor function and physiology was not easy because their actions on AMPA receptor physiology are not always direct and in some cases, it is the AMPA receptors which affect the function of these genes. For example, it is believed that 3-OST-3_A is important for brain morphogenesis and that AMPA receptor activity might help regulate the actions of 3-OST-3_A by inhibiting the release of HSPG (a substrate for 3-OST-3_A). In some cases, the role of these genes in AMPA receptor physiology is more straightforward and obvious. For example, RAB3B is believed to be involved in the exocytotic mechanism which delivers AMPA receptors to the cell surface. Ultimately, the

only way to prove conclusively the role of these genes in AMPA receptor physiology is to carry out knockdown or knockout studies of these genes in the wet lab.

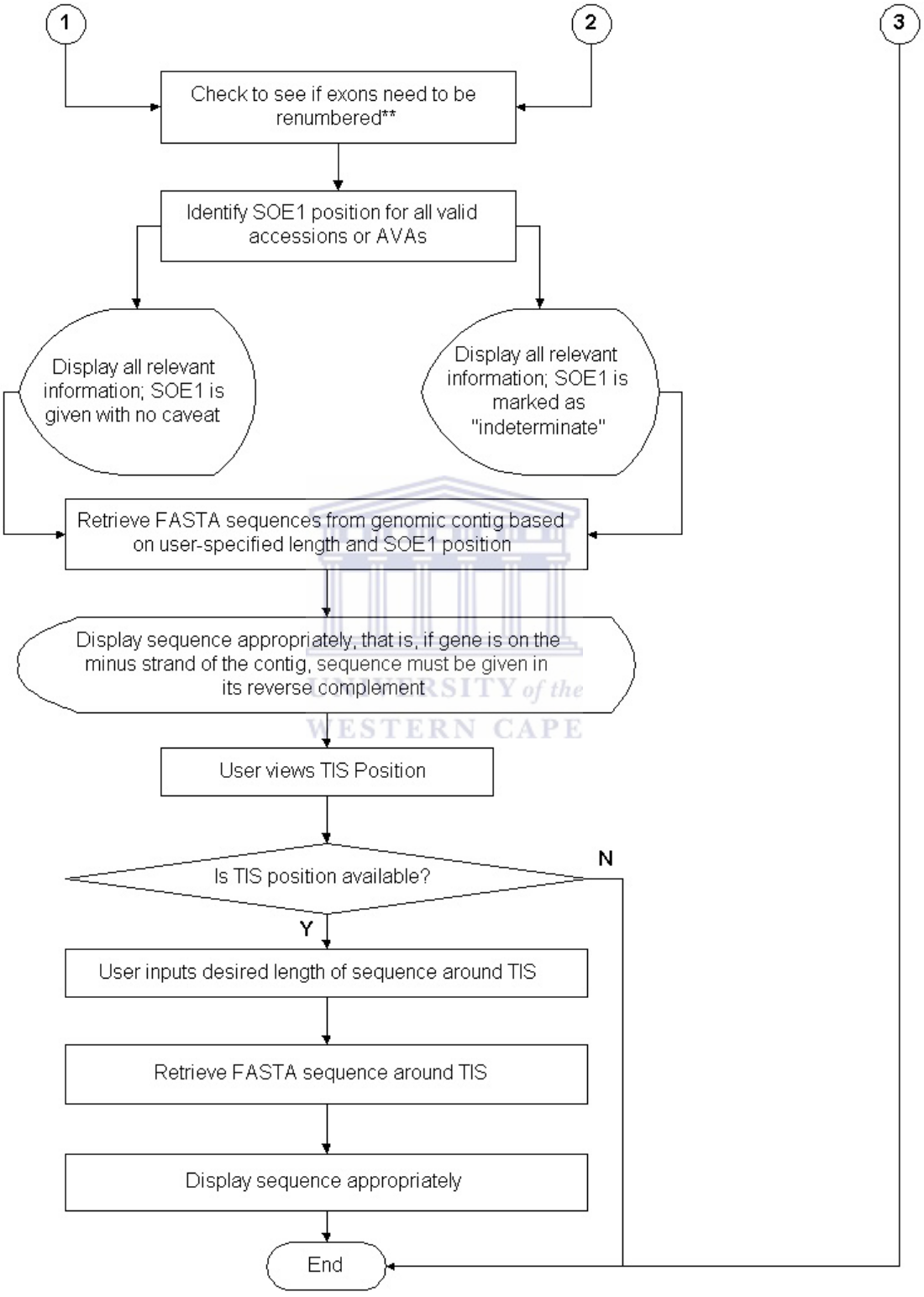
APPENDIX 1: FLOWCHARTS OF FIE2'S ALGORITHM

Key to Appendix 1: EV= LocusLink's Evidence Viewer page; CDS = protein coding sequence; SOE1 = start of exon 1; AVA = associated valid accession; ** = see respective flowchart for algorithm (given in sections A.2, A.3, or A.4); *** = see flowchart in section A.5 for algorithm

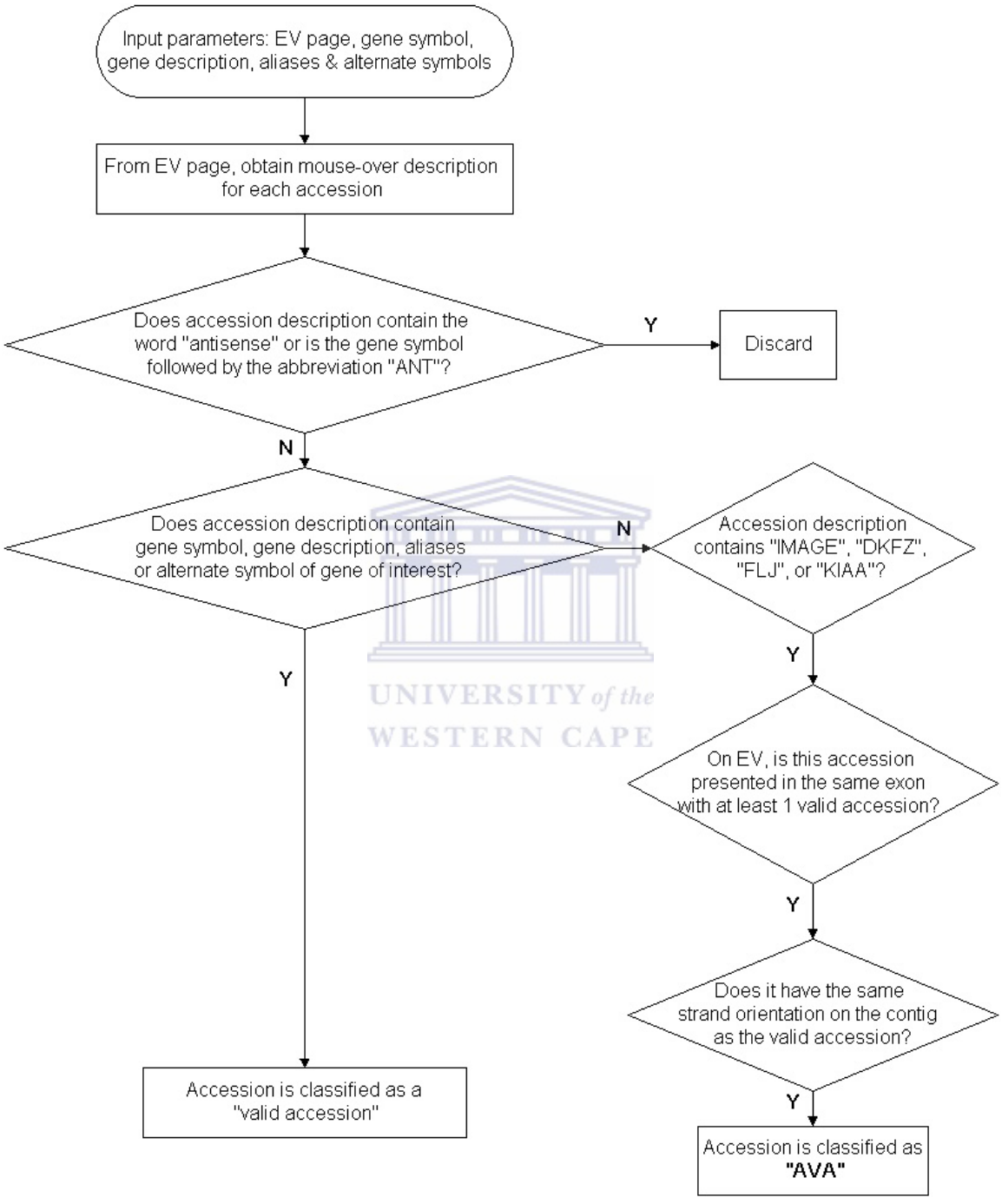
A.1.1 Overall algorithm for FIE2



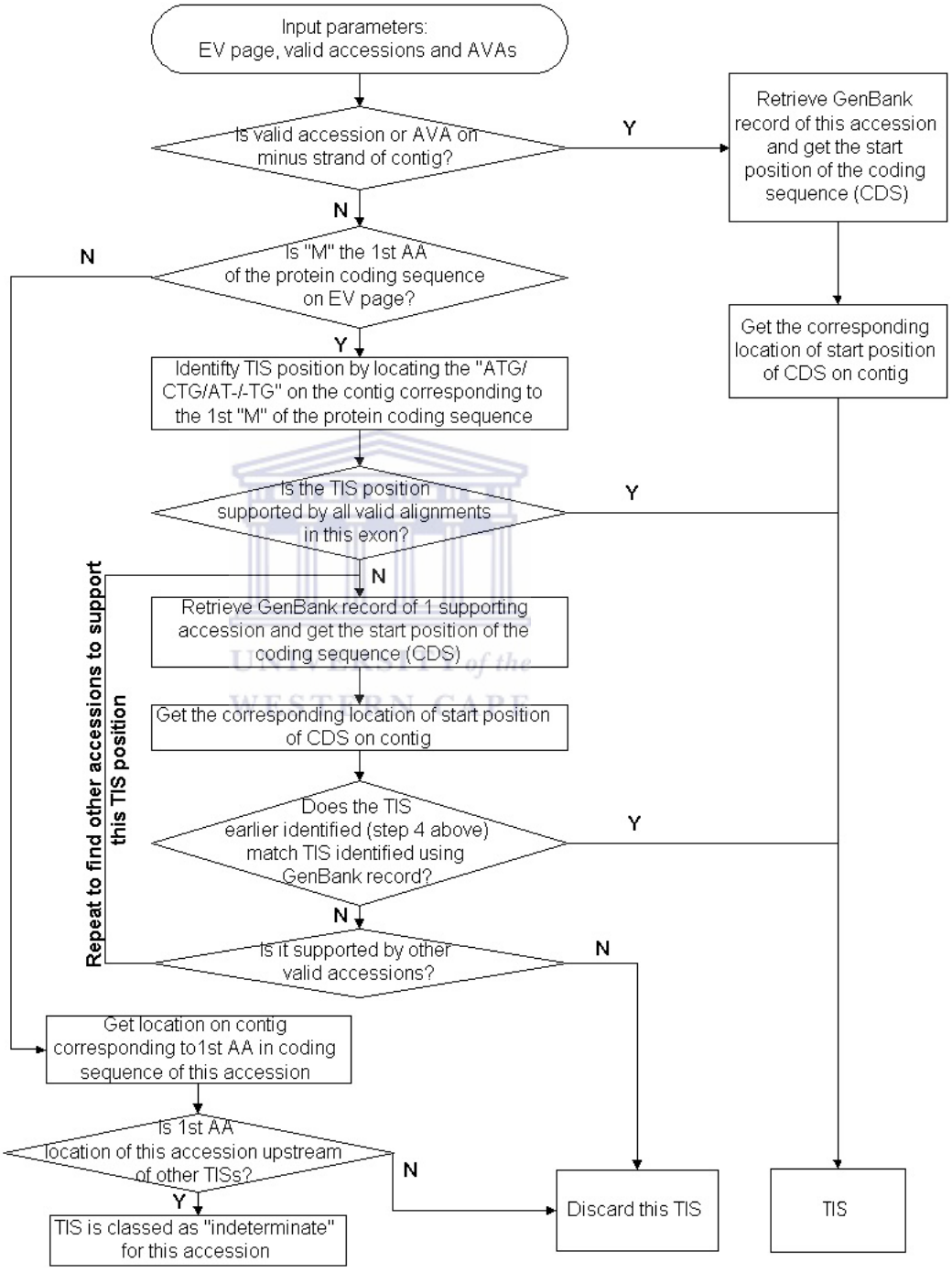
A.1.1 Overall algorithm for FIE2 (con't)



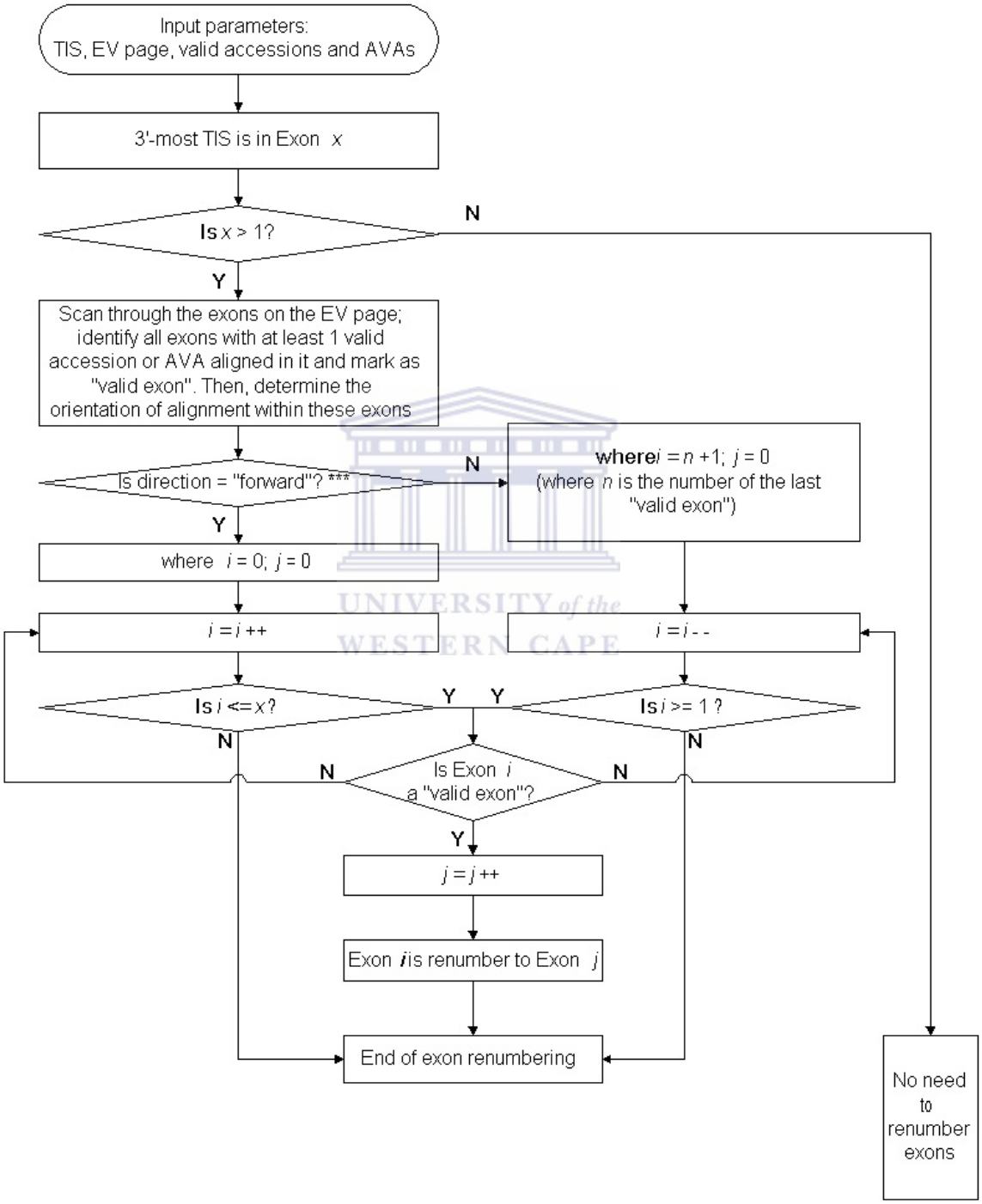
A.1.2 Identification of valid accessions



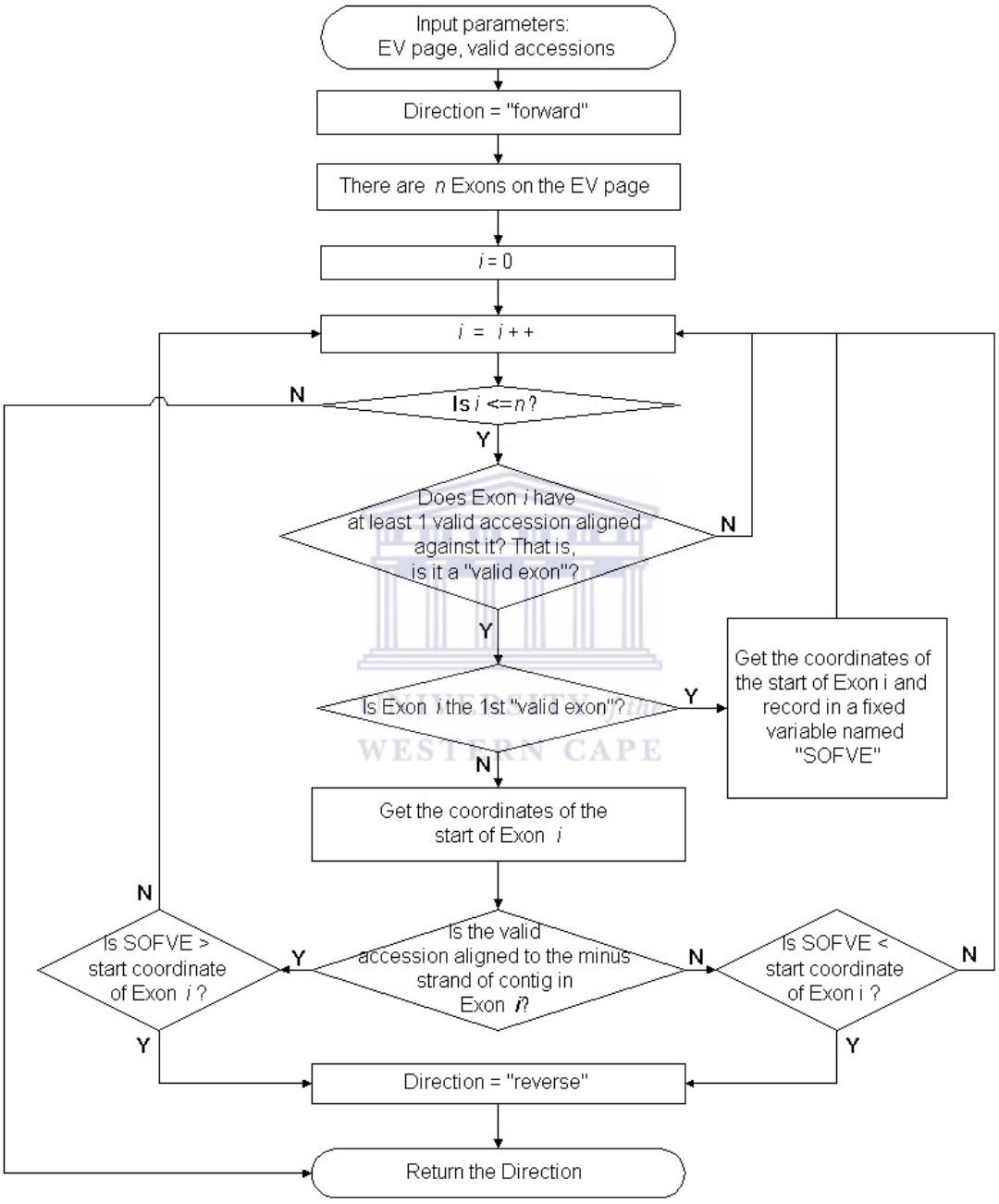
A.1.3 Locating the TIS position



A.1.4 Renumbering of exons



A.1.5 Determining the strand orientation of the gene on the genomic contig



APPENDIX 2: FIE2'S RETRIEVAL OF THE START POSITIONS OF GENES OF HUMAN CHROMOSOME 22 (RETRIEVED ON 21ST FEB. 2003)

No.	LocusID	Symbol	Description	Contig	SOE1 (position given with respect to the respective contig)	TIS (position given with respect to the respective contig)
1	50	ACO2	aconitase 2, mitochondrial	NT_011520.8	21079915 21110514 21139631	21079935
2	49	ACR	acrosin	NT_011526.4	313647	313664
3	135	ADORA2A	adenosine A2a receptor	NT_011520.8	4119995 4125484 4125765	4125765
4	157	ADRBK2	adrenergic, beta, receptor kinase 2	NT_011520.8	5257214 5257299 5257339 5257349	5257361
5	158	ADSL	adenylosuccinate lyase	NT_011520.8	19957297 19957317 19957318 19957347	19957347 19957422
6	162	AP1B1	adaptor-related protein complex 1, beta 1 subunit	NT_011520.8	9080919 9059609	9059582
7	23780	APOL2	apolipoprotein L, 2	NT_011520.8	15879559 15879248 15879180 15879112	15872767
8	80833	APOL3	apolipoprotein L, 3	NT_011520.8	15805784 15800536 15800370 15800362 15800344	15800498 15781415
9	80832	APOL4	apolipoprotein L, 4	NT_011520.8	15844438 15844348 15841436	15841641 15841339
10	80831	APOL5	apolipoprotein L, 5	NT_011520.8	15357478	15357478

11	80830	APOL6	apolipoprotein L, 6	NT_011520.8	15287983 15288019 15288021 15298389	15296032
12	55738	ARFGAP1	ADP-ribosylation factor GTPase activating protein 1	NT_011333.5	640796 640848 644463 646122 653773 654409	643564 644482 646961
13	23779	ARHGAP8	Rho GTPase activating protein 8	NT_011522.3	332218 332246 369752 380407 386045 386234 389269 407719 441618	381759 386872 441689
14	410	ARSA	arylsulfatase A	NT_011526.4	203576 203571 203531	203201
15	421	ARVCF	armadillo repeat gene deletes in velocardiofacial syndrome	NT_011519.9	3152263 3126289 3114688	3126271 3114519
16	468	ATF4	activating transcription factor 4 (tax-responsive enhancer element B67)	NT_011520.8	19131353 19131370 19131372 19131377 19132143	19132235
17	529	ATP6V1E1	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E isoform 1	NT_011519.9	1259516 1259468	1259382
18	613	BCR	breakpoint cluster region	NT_011520.8	2821339 2821494 2821641 2821650 2822047 2929265 2936199 2951431	2822090 2822255
19	637	BID	BH3 interacting	NT_011519.9	1405226	1380863

			domain death agonist		1405220 1405194 1404747 1404717 1380921	
20	638	BIK	BCL2-interacting killer (apoptosis-inducing)	NT_011520.8	22709444 22722709	22722716
21	23774	BRD1	bromodomain containing 1	NT_011525.4	924845 874405	924358
22	706	BZRP	benzodiazapine receptor (peripheral)	NT_011520.8	22750240 22750257 22758009	22757932 22759844
23	25776	C22orf2	chromosome 22 open reading frame 2	NT_011520.8	18267462 18267464	18278844
24	25807	C22orf3	chromosome 22 open reading frame 3	NT_011520.8	8960264 8957988	8957965
25	25771	C22orf4	chromosome 22 open reading frame 4	NT_011523.8	670680 670700 670725 670730 670735 945118 1082418	670835 786750
26	25829	C22orf5	chromosome 22 open reading frame 5	NT_011520.8	17883781 17841519 17834013 17832261	17858650
27	10369	CACNG2	calcium channel, voltage-dependent, gamma subunit 2	NT_011520.8	16313688 16313406	16313406
28	23466	CBX6	chromobox homolog 6	NT_011520.8	18483038 18483008 18483002	18482979
29	8318	CDC45L	CDC45 cell division cycle 45-like (S. cerevisiae)	NT_011519.9	2615318 2615325 2615345 2615349 2615358	2615373
30	51816	CECR1	cat eye syndrome chromosome region, candidate 1	NT_011519.9	842496 832283 832232	842284 832206

31	9620	CELSR1	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila)	NT_011523.8	445218 273428	445218
32	11200	CHEK2	CHK2 checkpoint homolog (S. pombe)	NT_011520.8	8437047 8434760 8427065 8427059	8427059 8411752
33	1120	CHKL	choline kinase-like	NT_011526.4	158394 158262 158253 158207	158210
34	7122	CLDN5	claudin 5 (transmembrane protein deleted in velocardiofacial syndrome)	NT_011519.9	2660741 2659804 2659776	2659659
35	8218	CLTCL1	clathrin, heavy polypeptide-like 1	NT_011519.9	2427089 2358215	2427045
36	1312	COMT	catechol-O-methyltransferase	NT_011519.9	3077349 3077351 3077374 3096676 3096695 3098024 3104146	3098004 3098154
37	1375	CPT1B	carnitine palmitoyltransferase 1B (muscle)	NT_011526.4	158394 158262 158253 158207 153996 153878	153344
38	1399	CRKL	v-crk sarcoma virus CT10 oncogene homolog (avian)-like	NT_011520.8	570689 570697 603994 604159 605829	571198
39	1413	CRYBA4	crystallin, beta A4	NT_011520.8	6314281	6314914
40	1414	CRYBB1	crystallin, beta B1	NT_011520.8	6310391 6310344	6308636
41	1415	CRYBB2	crystallin, beta B2	NT_011520.8	4913738	4913750
42	1417	CRYBB3	crystallin, beta B3	NT_011520.8	4892178	4893717
43	1439	CSF2RB	colony stimulating	NT_011520.8	16533007	16533035

			factor 2 receptor, beta, low-affinity (granulocyte-macrophage)			
44	1454	CSNK1E	casein kinase 1, epsilon	NT_011520.8	18009311 17928874 17928169 17928009 17928003 17924959	17924947
45	1652	DDT	D-dopachrome tautomerase	NT_011520.8	3618417 3613035 3613026 3613010	3613001
46	10521	DDX17	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 17, 72kDa	NT_011520.8	18117097 18117083 18117072 18116789 18097229	18116789
47	9993	DGCR2	DiGeorge syndrome critical region gene 2	NT_011519.9	2257810 2257800 2257799 2257759	2257615
48	8214	DGCR6	DiGeorge syndrome critical region gene 6	NT_011519.9	2041438 2041689 2041699 2041736 2041765 2041817	2041785 2046334
49	1727	DIA1	diaphorase (NADH) (cytochrome b-5 reductase)	NT_011520.8	22248064 22248036 22245705	22248008 22235491
50	11144	DMC1	DMC1 dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (yeast)	NT_011520.8	18180973 18180851 18179077	18179045
51	10126	DNAL4	dynein, axonemal, light polypeptide 4	NT_011520.8	18404938 18404912	18393520
52	4733	DRG1	developmentally regulated GTP binding protein 1	NT_011520.8	11091906 11091911 11091914	11091972
53	1890	ECGF1	endothelial cell	NT_011526.4	105455	105138

			growth factor 1 (platelet-derived)		105450 105441 105425	
54	8664	EIF3S7	eukaryotic translation initiation factor 3, subunit 7 zeta, 66/67kDa	NT_011520.8	16140268 16139981 16139968	16136953
55	2033	EP300	E1A binding protein p300	NT_011520.8	20702574	20703793
56	2192	FBLN1	fibulin 1	NT_011522.3	1158044 1158137 1253995	1158147
57	25793	FBXO7	F-box only protein 7	NT_011520.8	12167011 12167288 12171269 12184075	12167291
58	2547	G22P1	thyroid autoantigen 70kDa (Ku antigen)	NT_011520.8	21232083 21232110 21232138 21246958	21232793
59	8484	GALR3	galanin receptor 3	NT_011520.8	17434174	17434199
60	10634	GAS2L1	growth arrest-specific 2 like 1	NT_011520.8	8999347 8999381 8999395 8999398	9000446
61	23464	GCAT	glycine C-acetyltransferase (2-amino-3-ketobutyrate coenzyme A ligase)	NT_011520.8	17418740 17418757	17418760
62	26088	GGA1	golgi associated, gamma adaptin ear containing, ARF binding protein 1	NT_011520.8	17219288 17219626 17219627 17219638 17219648 17219649 17219832 17231640 17234621	17219653
63	2687	GGTLA1	gamma-glutamyltransferase-like activity 1	NT_011520.8	3937492 3937420	3937085
64	2781	GNAZ	guanine nucleotide binding protein (G	NT_011520.8	2711482 2711611	2736825

			protein), alpha z polypeptide		2736821 2736825	
65	54584	GNB1L	guanine nucleotide binding protein (G protein), beta polypeptide 1-like	NT_011519.9	2990342 2990307 2990299 2990266 2990243	2956758
66	2812	GP1BB	glycoprotein Ib (platelet), beta polypeptide	NT_011519.9	2858775 2858946	2858973
67	2847	GPR24	G protein-coupled receptor 24	NT_011520.8	20290021	20290234
68	9402	GRAP2	GRB2-related adaptor protein 2	NT_011520.8	19511870 19511910 19511926 19511936 19511941 19511965 19511973 19511979 19557605 19557640 19557881 19557895	19557895
69	2952	GSTT1	glutathione S-transferase theta 1	NT_011520.8	3680669 3680658 3680629	3680629
70	2953	GSTT2	glutathione S-transferase theta 2	NT_011520.8	3618722 3618732 3618787	3618787
71	9567	GTPBP1	GTP binding protein 1	NT_011520.8	18316753 18316847 18338919 18339045 18341769 18343813	18319697
72	51512	GTSE1	G-2 and S-phase expressed 1	NT_011523.8	204963 205436 205458	205506
73	3005	H1F0	H1 histone family, member 0	NT_011520.8	17415959 17416021 17416063	17416337
74	7290	HIRA	HIR histone cell cycle regulation defective homolog	NT_011519.9	2567128 2567100 2566941	2566880 2543965

			A (S. cerevisiae)		2546182	
75	3162	HMOX1	heme oxygenase (decycling) 1	NT_011520.8	15073387 15073401	15073467
76	23765	IL17R	interleukin 17 receptor	NT_011519.9	717906 717953 717995	718030
77	3560	IL2RB	interleukin 2 receptor, beta	NT_011520.8	16755030	16754997
78	3761	KCNJ4	potassium inwardly-rectifying channel, subfamily J, member 4	NT_011520.8	18065987 18065963 18065872 18054817	18038921
79	27093	KCNMB3L	potassium large conductance calcium-activated channel, subfamily M, beta member 3-like	NT_011519.9	225757 225740 225737	225497
80	11015	KDELR3	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3	NT_011520.8	18078867 18078874	18079023
81	9215	LARGE	like-glycosyltransferase	NT_011520.8	13612710 13612708 13612626	13453763
82	3956	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1)	NT_011520.8	17286428 17286446	17286495
83	3957	LGALS2	lectin, galactoside-binding, soluble, 2 (galectin 2)	NT_011520.8	17190723 17190718	17190697
84	3976	LIF	leukemia inhibitory factor (cholinergic differentiation factor)	NT_011520.8	9939050 9939030	9938986
85	3985	LIMK2	LIM domain kinase 2	NT_011520.8	10904552 10904567 10940690 10940775 10970150	10904697 10941005
86	8216	LZTR1	leucine-zipper-like transcriptional regulator, 1	NT_011520.8	635630 635639 648327	644939

87	23764	MAFF	v-maf musculoaponeurotic fibrosarcoma oncogene homolog F (avian)	NT_011520.8	17812783 17812803 17812851 17812893 17824327	17824646
88	5594	MAPK1	mitogen-activated protein kinase 1	NT_011520.8	1520917 1520876 1520862 1520846 1520662	1520677 1520641
89	5600	MAPK11	mitogen-activated protein kinase 11	NT_019197.3	283797 283770 283723 283718 283716 283712 280838	283712
90	6300	MAPK12	mitogen-activated protein kinase 12	NT_019197.3	275238 275080 274962 270527	274929
91	23542	MAPK8IP2	mitogen-activated protein kinase 8 interacting protein 2	NT_011526.4	176131 176221 178562	176248 178562
92	4151	MB	myoglobin	NT_011520.8	15256933 15255319	15256863 15250642
93	4174	MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	NT_011520.8	15092428 15092464 15092724	15092732
94	4242	MFNG	manic fringe homolog (Drosophila)	NT_011520.8	17097170	17097000
95	4248	MGAT3	mannosyl (beta-1,4-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase	NT_011520.8	19068109 19096488	19098143
96	4282	MIF	macrophage migration inhibitory factor (glycosylation-	NT_011520.8	3532963 3532978 3532985 3532996	3533060

			inhibiting factor)		3533009 3533028 3533061 3533086	
97	23786	BCL2L13	BCL2-like 13 (apoptosis facilitator)	NT_011519.9	1259663 1269488 1269571 1286459 1319731 1319733 1357996 1358146	1286459 1319801
98	57591	MKL1	megakaryoblastic leukemia (translocation) 1	NT_011520.8	20247475 20247432 20034370 20029331 20023676	20074015 20031753
99	4320	MMP11	matrix metalloproteinase 11 (stromelysin 3)	NT_011520.8	3411434	3411456
100	4330	MN1	meningioma (disrupted in balanced translocation) 1	NT_011520.8	7493836 7443303	7492950 7492881
101	4357	MPST	mercaptopyruvate sulfurtransferase	NT_011520.8	16630576 16630607 16633863 16635018	16635042
102	11135	CDC42EP1	CDC42 effector protein (Rho GTPase binding) 1	NT_011520.8	17171285 17171316	17177142
103	8897	MTMR3	myotubularin related protein 3	NT_011520.8	9575501 9575541 9670733	9663351
104	4627	MYH9	myosin, heavy polypeptide 9, non-muscle	NT_011520.8	16027530 15988859 15988840 15981407 15925317 15922830 15921924	15988840
105	4668	NAGA	N-acetylgalactosaminidase, alpha-	NT_011520.8	21681557 21681166	21681085
106	4689	NCF4	neutrophil cytosolic	NT_011520.8	16471815	16471999

			factor 4, 40kDa		16471843 16471940 16471963	
107	4700	NDUFA6	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6, 14kDa	NT_011520.8	21701533	21701532
108	4744	NEFH	neurofilament, heavy polypeptide 200kDa	NT_011520.8	9172566 9172599	9172599
109	4771	NF2	neurofibromin 2 (bilateral acoustic neuroma)	NT_011520.8	9295897 9295956 9295984 9296050 9296105 9296272 9296322 9388014 9389160 9389453 9389546 9390232 9390332	9296322
110	4809	NHP2L1	NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae)	NT_011520.8	21299660 21299635 21293240 21291153 21291152	21299584 21293146
111	8508	NIPSNAP1	nipsnap homolog 1 (C. elegans)	NT_011520.8	9273660 9273410 9273406 9262850	9273405
112	64976	MRPL40	mitochondrial ribosomal protein L40	NT_011519.9	2567917 2567947 2567949 2567959	2567959
113	23467	NPTXR	neuronal pentraxin receptor	NT_011520.8	18454801 18454115 18433695	18454647
114	10762	NUP50	nucleoporin 50kDa	NT_011522.3	819007 819228 819304 819444 819808	823340 826777

					819833 832861 839762	
115	23762	OSBP2	oxysterol binding protein 2	NT_011520.8	10387095 10387163 10387201 10456542 10579715 10583239	10387199 10387313
116	5008	OSM	oncostatin M	NT_011520.8	9959131	9959090
117	9127	P2RXL1	purinergic receptor P2X-like 1, orphan receptor	NT_011520.8	668398 668424 668429	668443
118	11252	PACSIN2	protein kinase C and casein kinase substrate in neurons 2	NT_011520.8	22613838 22545735 22510850 22477766	22510773
119	29780	PARVB	parvin, beta	NT_011521.1	503234 503272 503298	503281 572896
120	5155	PDGFB	platelet-derived growth factor beta polypeptide (simian sarcoma viral (v-sis) oncogene homolog)	NT_011520.8	18855844 18855774 18855064 18854869 18851698 18846663 18836052 18835851	18854752 18851661
121	23481	PES1	pescadillo homolog 1, containing BRCT domain (zebrafish)	NT_011520.8	10299274 10284196 10284175 10284171	10284122
122	27124	PIB5PA	phosphatidylinositol (4,5) bisphosphate 5-phosphatase, A	NT_011520.8	10815254 10815282	10818129
123	5297	PIK4CA	phosphatidylinositol 4-kinase, catalytic, alpha polypeptide	NT_011520.8	492022 395916 387959 380649	492022 387842
124	23760	PITPNB	phosphotidylinositol transfer protein, beta	NT_011520.8	7611568 7611561	7611529
125	8563	C22orf19	chromosome 22 open reading frame	NT_011520.8	9246037 9246020	9241470

			19		9246011 9235819	
126	10343	PKDREJ	polycystic kidney disease (polycystin) and REJ (sperm receptor for egg jelly homolog, sea urchin)-like	NT_011523.8	171370	171370
127	8398	PLA2G6	phospholipase A2, group VI (cytosolic, calcium-independent)	NT_011520.8	17792620 17792548 17780218 17753930 17730106	17780218
128	5372	PMM1	phosphomannomutase 1	NT_011520.8	21200655 21200608 21200597 21191477	21200593
129	5413	PNUTL1	peanut-like 1 (Drosophila)	NT_011519.9	2849905 2849947 2853861 2857120 2858348 2858775	2849992 2854141
130	5435	POLR2F	polymerase (RNA) II (DNA directed) polypeptide F	NT_011520.8	17564480 17564491 17564503	17564580
131	5465	PPARA	peroxisome proliferative activated receptor, alpha	NT_011523.8	59938 84801 84812	106432
132	23759	PPIL2	peptidylprolyl isomerase (cyclophilin)-like 2	NT_011520.8	1319223 1319257 1319258 1319275 1347897	1319336
133	5625	PRODH	proline dehydrogenase (oxidase) 1	NT_011519.9	2071962 2057814 2055041 2053823	2071696
134	27128	PSCD4	pleckstrin homology, Sec7 and coiled/coil domains 4	NT_011520.8	16893209 16893280 16893341 16911797	16893396 16914112
135	5816	PVALB	parvalbumin	NT_011520.8	16430308	16427834

136	9609	RAB36	RAB36, member RAS oncogene family	NT_011520.8	2786455 2786485 2803692	2786495
137	11158	RABL2B	RAB, member of RAS oncogene family-like 2B	NT_011526.4	359057 359048 359036 359027 348975 344551	357721
138	5880	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	NT_011520.8	16855090 16855080 16854973 16852482 16842173 16836820	16854973 16836802
139	5902	RANBP1	RAN binding protein 1	NT_011519.9	3251415 3252978 3253025 3259616	3253127
140	5905	RANGAP1	Ran GTPase activating protein 1	NT_011520.8	20896954 20891832	20891832
141	11020	RABL4	RAB, member of RAS oncogene family-like 4	NT_011520.8	16386942 16386899 16372397	16386536
142	23543	RBM9	RNA binding motif protein 9	NT_011520.8	15668108 15667957 15480013 15479987 15388238	15667852 15479824
143	9978	RBX1	ring-box 1	NT_011520.8	20562169 20562182 20562187 20562224 20576989	20562187
144	5988	RFPL1	ret finger protein-like 1	NT_011520.8	9130922	9131218
145	10739	RFPL2	ret finger protein-like 2	NT_011520.8	11895765 11894034 11885767	11885475
146	10738	RFPL3	ret finger protein-like 3	NT_011520.8	12047596 12050155 12050440	12050447
147	6122	RPL3	ribosomal protein L3	NT_011520.8	18931175 18930412	18930386

					18930400 18929382 18929368	
148	10633	RRP22	RAS-related on chromosome 22	NT_011520.8	9008095 9007906	9007585
149	27156	RTDR1	rhabdoid tumor deletion region gene 1	NT_011520.8	2783183 2783151	2781549
150	9997	SCO2	SCO cytochrome oxidase deficient homolog 2 (yeast)	NT_011526.4	100972	99840
151	23753	SDF2L1	stromal cell-derived factor 2-like 1	NT_011520.8	1295548 1295553	1295573
152	55964	SEPT3	septin 3	NT_011520.8	21587780 21587838	21592462
153	3053	SERPIND1	serine (or cysteine) proteinase inhibitor, clade D (heparin cofactor), member 1	NT_011520.8	427403 432585 432656	432601
154	23544	SEZ6L	seizure related 6 homolog (mouse)-like	NT_011520.8	5984570	5985312
155	10291	SF3A1	splicing factor 3a, subunit 1, 120kDa	NT_011520.8	10049207 10049179 10033095	10049083
156	23616	SH3BP1	SH3-domain binding protein 1	NT_011520.8	17251918 17252163	17253364 17264654
157	23539	SLC16A8	solute carrier 16 (monocarboxylic acid transporters), member 8	NT_011520.8	17693934	17693665
158	6576	SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	NT_011519.9	2314181 2314156 2314136 2314133 2314030	2314082
159	10478	SLC25A17	solute carrier family 25 (mitochondrial carrier; peroxisomal membrane protein, 34kDa), member 17	NT_011520.8	20430157 20430147 20430125 20388161	20430036
160	83733	SLC25A18	solute carrier family	NT_011519.9	1194230	1210692

			25, (mitochondrial carrier), member 18		1194231 1194336	
161	6523	SLC5A1	solute carrier family 5 (sodium/glucose cotransporter), member 1	NT_011520.8	11735560	11735570
162	6527	SLC5A4	solute carrier family 5 (low affinity glucose cotransporter), member 4	NT_011520.8	11947620 11944154	11947617
163	6545	SLC7A4	solute carrier family 7 (cationic amino acid transporter, y+ system), member 4	NT_011520.8	685796 685090	685050
164	6598	SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	NT_011520.8	3425548 3425583 3425608 3425686 3425708 3472157	3425755
165	27127	SMC1L2	SMC1 structural maintenance of chromosomes 1-like 2 (yeast)	NT_011522.3	1068729	1068729
166	9342	SNAP29	synaptosomal-associated protein, 29kDa	NT_011520.8	512283 512316 512323 536739	512390
167	6634	SNRPD3	small nuclear ribonucleoprotein D3 polypeptide 18kDa	NT_011520.8	4248313 4248340 4248350	4250021
168	6663	SOX10	SRY (sex determining region Y)-box 10	NT_011520.8	17598214 17595324 17595161 17594660	17594576
169	6721	SREBF2	sterol regulatory element binding transcription factor 2	NT_011520.8	21443942 21496384	21444059 21511158
170	6753	SSTR3	somatostatin receptor 3	NT_011520.8	16818627	16818627

171	6767	ST13	suppression of tumorigenicity 13 (colon carcinoma) (Hsp70 interacting protein)	NT_011520.8	20467471 20467403 20467400 20437175 20436034	20467328 20435610
172	25830	SULT4A1	sulfotransferase family 4A, member 1	NT_011521.1	341459 341397 341396 341352 341329 341323 305974	341323
173	8224	SYN3	synapsin III	NT_011520.8	12698983	12698947
174	9145	SYNGR1	synaptogyrin 1	NT_011520.8	18960773 18960800 18974959	18960800 18974984
175	10454	MAP3K7IP1	mitogen-activated protein kinase kinase kinase 7 interacting protein 1	NT_011520.8	19010572 19010579	19010592
176	6899	TBX1	T-box 1	NT_011519.9	2892106	2895047
177	6948	TCN2	transcobalamin II; macrocytic anemia	NT_011520.8	10299463 10299493 10299494 10299524 10299568	10299621
178	7008	TEF	thyrotrophic embryonic factor	NT_011520.8	20978243 20992749 20992798 20992909 21007508 21008623	20992833 20992959
179	7078	TIMP3	tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory)	NT_011520.8	12493108 12494008 12494011 12494045 12494218 12541739 12551763 12552306 12553293	12494288 12552725
180	10766	TOB2	transducer of ERBB2, 2	NT_011520.8	21048347	21048133
181	10040	TOM1L1	target of myb1-like	NT_010783.11	7402958	7403072

			1 (chicken)		7403049 7403063	
182	8940	TOP3B	topoisomerase (DNA) III beta	NT_011520.8	1636134 1636094 1629126 1629101 1620969	1629028
183	8459	TPST2	tyrosylprotein sulfotransferase 2	NT_011520.8	6282406 6257687 6234037	6233949
184	10587	TXNRD2	thioredoxin reductase 2	NT_011519.9	3077555 3077371 3077366 3073742 3068116 3066648 3051929 3019907 3016750	3077366 3067932 3054508 3033474
185	7263	TST	thiosulfate sulfurtransferase (rhodanese)	NT_011520.8	16630273 16630256 16630253 16630243	16629558
186	25809	TTLL1	tubulin tyrosine ligase-like 1	NT_011520.8	22688121 22688108 22688097 22688081 22688030 22688022 22674279	22674279
187	51807	TUBA8	tubulin, alpha 8	NT_011519.9	1741359 1741465	1741538
188	25828	TXN2	thioredoxin 2	NT_011520.8	16092472 16092433 16091669	16091669
189	7332	UBE2L3	ubiquitin-conjugating enzyme E2L 3	NT_011520.8	1220966 1246097	1220981
190	7353	UFD1L	ubiquitin fusion degradation 1-like	NT_011519.9	2614567 2614546 2607110	2614489
191	7380	UPK3	uroplakin 3	NT_011522.3	940149 941147	940176
192	7441	VPREB1	pre-B lymphocyte	NT_011520.8	1898147	1898173

			gene 1			
193	29802	VPREB3	pre-B lymphocyte gene 3	NT_011520.8	3393055 3392992	3392950
194	7494	XBP1	X-box binding protein 1	NT_011520.8	8492910 8492894 8492867 8488070	8492862
195	7533	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	NT_011520.8	11636830 11636838	11637021
196	23598	ZNF278	zinc finger protein 278	NT_011520.8	11038551 11038520 11038098 11037899	11037890
197	7625	ZNF74	zinc finger protein 74 (Cos52)	NT_011520.8	50656 50672 50983 51818	57193
198	84700	MYO18B	myosin XVIIIB	NT_011520.8	5434464 5434473 5462347 5684707 5719368	5453413 5462379 5695627
199	27341	CGI-96	CGI-96 protein	NT_011520.8	22118788 22118495 22116822 22115608	22118614
200	27350	APOBEC3C	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3C	NT_011520.8	18625046 18625049 18625064	18625152
201	84844	PHF5A	PHD finger protein 5A	NT_011520.8	21079475	21079441
202	27352	DJ1042K10.2	hypothetical protein DJ1042K10.2	NT_011520.8	20011484	20011595
203	51493	HSPC117	hypothetical protein HSPC117	NT_011520.8	12104536 12104524 12104519 12104513	12104444
204	27351	D15Wsu75e	DNA segment, Chr 15, Wayne State	NT_011520.8	21231795	21231627

			University 75, expressed			
205	11078	HRIHFB2122	Tara-like protein	NT_011520.8	17357032 17357103 17368674 17368751 17376471	17357207
206	60489	APOBEC3G	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G	NT_011520.8	18687794 18687920 18688092 18688109 18690451 18691424	18688151
207	9582	APOBEC3B	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	NT_011520.8	18593189 18593214 18597130 18602856	18593243 18596868 18600253
208	83746	L3MBTL2	l(3)mbt-like 2 (Drosophila)	NT_011520.8	20816070 20816100 20816108 20816112	20816151
209	6305	SBF1	SET binding factor 1	NT_011526.4	50464 41857 40670 (Indeterminate 1)	Unavailable
210	6942	TCF20	transcription factor 20 (AR1)	NT_011520.8	21810701 21759675 (Indeterminate 1)	Unavailable
211	9701	KIAA0685	KIAA0685 gene product	NT_011526.4	6601 (Indeterminate 1)	Unavailable
212	51586	PCQAP	PC2 (positive cofactor 2, multiprotein complex) glutamine/Q-rich-associated protein	NT_011520.8	164056 164067 164075 164104 193582 211401 223073 231282 238370 239582 (Indeterminate 2)	Indeterminate 164144 193592
213	23761	PISD	phosphatidylserine decarboxylase	NT_011520.8	11354503 11323112 (Indeterminate 2)	Indeterminate 11318103
214	27254	PIPPIN	ortholog of rat	NT_011520.8	21171913	Indeterminate

			pippin		21182734 (Indeterminate 2)	21182754
215	10042	HMG2L1	high-mobility group protein 2-like 1	NT_011520.8	14976255 (Indeterminate 3)	Unavailable
216	10737	RFPL3S	ret finger protein-like 3 antisense	NT_011520.8	11880697 (Indeterminate 3)	Unavailable
217	94009	SERHL	serine hydrolase-like	NT_011520.8	22099272 (Indeterminate 3)	Unavailable
218	8542	APOL1	apolipoprotein L, 1		No EV page	
219	758	C22orf1	chromosome 22 open reading frame 1		No EV page	
220	23492	CBX7	chromobox homolog 7		No EV page	
221	2130	EWSR1	Ewing sarcoma breakpoint region 1		No EV page	
222	2678	GGT1	gamma-glutamyltransferase 1		No EV page	
223	23532	PRAME	preferentially expressed antigen in melanoma		No EV page	
224	9463	PRKCABP	protein kinase C, alpha binding protein		No EV page	
225	10740	RFPL1S	ret finger protein-like 1 antisense		No EV page	
226	9490	SCA10	spinocerebellar ataxia 10		No EV page	
227	23541	SEC14L2	SEC14-like 2 (S. cerevisiae)		No EV page	
228	6525	SMTN	smoothelin		No EV page	
229	23752	STK22A	serine/threonine kinase 22A (spermiogenesis associated)		No EV page	
230	23748	ZNF279	zinc finger protein 279		No EV page	

APPENDIX 3: A COMPARISON OF FIE2'S EXTRACTIONS AGAINST SANGER'S ANNOTATIONS

Tables comparing FIE2 extracted SOE1 positions with annotated gene start positions given by Sanger

* indicates that at least one of the extracted SOE1 position is the same as the annotated gene start position from Sanger.

i) 5'-most SOE1 position extracted by FIE2 is identical to Sanger's annotations

Gene Symbol	SOE1		Difference	Strand	
	FIE2	Sanger		FIE 2	Sanger
TOM1	32391891 32391907	32391891	0	+	+
PKDREJ	43225483	43225483	0	-	-
TXNRD2	16869902 16869718 16869713 16866089 16858995 16844276	16869902	0	-	-
RFPL3	29449901	29449901	0	+	+
RFPL2	29295511 29285513	29295511	0	-	-
CRKL	17970435	17970435	0	+	+
CRYBA4	23714027	23714027	0	+	+
CRYBB3	22291924	22291924	0	+	+
DDT	21018163 21012781 21012772 21012756	21018163	0	-	-
EP300	38102320	38102320	0	+	+
NPTXR	35854547 35853861 35833441	35854547	0	-	-
MAFF	35212529 35212549 35212597 35212639	35212529	0	+	+
C22orf5	35283527 35232007	35283527	0	-	-

PSCD4	34292955 34293026 34293087 34311543	34292955	0	+	+
GPR24	37689767 37691254	37689767	0	+	+
GSCL	16078039	16078039	0	-	-
VPREB3	20792801 20792738	20792801	0	-	-
IGLL1	20618642 20618582	20618642	0	-	-
ARVCF	16944610	16944610	0	-	-
MN1	24893582 24843049	24893582	0	-	-
NF2	26695643 26695702 26695730 26695796 26695851 26696018 26696068 26789199 26789978 26790078	26695643	0	+	+
ACR	47680667	47680667	0	+	+
PMM1	38600401 38600354 38600343	38600401	0	-	-
GNB1L	16782689 16782654 16782590	16782689	0	-	-
PRODH	15864309 15858942 15847354 15846170	15864309	0	-	-
SMTN	28173353 28183383	28173353	0	+	+
TBX1	16684453	16684453	0	+	+
HIRA	16359447 16359288 16338529	16359447	0	-	-
VPREB1	19297893	19297893	0	+	+
PLA2G6	35192294 35179964	35192294	0	-	-

	35153676				
GALR3	34833920	34833920	0	+	+
NIPSNAP1	26673406 26673152	26673406	0	-	-
CACNA1I	36581288	36581288	0	+	+
P2RXL1	18068144 18068170	18068144	0	+	+
SNAP29	17912029 17912062 17912065 17936485	17912029	0	+	+
GTPBP1	35716499 35716593	35716499	0	+	+
RAB36	20186201 20186231	20186201	0	+	+
CELSR1	43499331 43324282	43499331	0	-	-

ii) 5'-most SOE1 extracted by FIE2 is upstream of Sanger's annotated position

Gene Symbol	SOE1		Difference	Strand	
	FIE2	Sanger		FIE 2	Sanger
LIMK2	28304298 28340436 28340521	28304299	1	+	+
BID	15197573 15197541 15173268	15197572	1	-	-
LARGE	31012456 31012454 31012372	31012455	1	-	-
NEFH	26572312 26572345	26572314	2	+	+
SLC25A1	16106528 16106503 16106483 16106480	16106526	2	-	-
ADRBK2	22657045 22657095	22657050	5	+	+
IL2RB	34160493 34160476	34160487	6	-	-
DGCR6	15834036 15834046 15834083 15834164	15834045	9	+	+
CLTCL1	16219436 16150562	16219425	11	-	-
MMP11	20811180	20811193	13	+	+
LIF	27338796 27338776	27338776	20*	-	-
MGAT3	36467855 36496234	36467879	24	+	+
TTLL1	40087867 40087854 40087776 40087768 40074025	40087840	27	-	-
GRAP2	36911616 36911656 36911672 36911682 36911687	36911646	30	+	+

	36911711 36911719 36911725 36957351 36957386 36957627 36957641				
MAPK11	47260313 47260286 47260239 47260234 47260232 47260228 47257354	47260278	35	-	-
MAPK1	18920663 18920622 18920592 18920408	18920612	51	-	-
ZNF278	28438297 28438266 28437844 28437645	28438236	61	-	-
SULT4A1	40875147 40875085 40875084 40875040 40875017 40875011	40875084	63*	-	-
GGTLA1	21337238 21337166	21337166	72*	-	-
BCR	20221085 20221240 20221387 20221396 20221793 20329011 20335945	20221240	155*	+	+
TST	34030019 34030002 34029989	34029744	275	-	-
PDGFB	36255590 36255520 36254810 36254615 36251444	36254615	975*	-	-

	36246409 36235798 36235597				
ECGF1	47474271	47472278	1993	-	-
MAPK8IP2	47543151 47545582	47545582	2431*	+	+
MPST	34030353 34033609 34034764	34034717	4364	+	+
TOP3B	19035840 19028872 19028847	19029034	6806	-	-
RBM9	33067854 33067703 32879759 32879733	32879943	187911	-	-

iii) 5'-most SOE1 extracted by FIE2 is downstream of Sanger's annotated position

Gene Symbol	SOE1		Difference	Strand	
	FIE2	Sanger		FIE 2	Sanger
TUBA8	15533706 15533812	15533704	2	+	+
MAPK12	47251754 47251596 47251478	47251756	2	-	-
BZRP	40149986 40150003 40157755	40149984	2	+	+
RAC2	34254836 34254719 34241919	34254839	3	-	-
SLC7A4	18085542 18084836	18085545	3	-	-
XBP1	25892656 25892640 25892613	25892660	4	-	-
SMARCB1	20825294 20825329 20825354 20825432 20825454 20871903	20825289	5	+	+

UPK3	42316192 42317190	42316187	5	+	+
POLR2F	34964223 34964249	34964215	8	+	+
KDELR3	35478613 35478620	35478604	9	+	+
SLC5A4	29347366 29343900	29347375	9	-	-
GSTT2	21018468 21018479 21018533	21018458	10	+	+
OSM	27358877	27358887	10	-	-
MCM5	32492174 32492210 32492470	32492162	12	+	+
COMT	16869696 16869721 16889023 16889042 16890371 16896493	16869683	13	+	+
PPIL2	18719004 18719021	18718991	13	+	+
CDC45L	16407665 16407672 16407692 16407696 16407705	16407650	15	+	+
SLC25A17	37829903 37829871	37829919	16	-	-
SYNGR1	36360519 36360546 36374705	36360503	16	+	+
GCAT	34818486 34818503	34818469	17	+	+
ADSL	37357043 37357064 37357093	37357025	18	+	+
ZNF74	17450402 17450418 17451564	17450384	18	+	+
DNAL4	35804684 35804658	35804703	19	-	-
PRAME	19600374	19600393	19	-	-

	19600114 19589354				
ACO2	38479661 38510260 38539377	38479641	20	+	+
PVALB	33830054	33830074	20	-	-
ARSA	47570596 47570591	47570617	21	-	-
RBX1	37961914 37961928 37961933 37961970	37961893	21	+	+
DIA1	39647810 39647782 39645451	39647836	26	-	-
DDX17	35516843 35516829 35516818 35516535 35496975	35516872	29	-	-
PACSIN2	40013584 39945481 39910596 39877512	40013613	29	-	-
ATP6V1E1	15051863 15051815	15051894	31	-	-
RPL3	36330158 36330146 36330145 36329128 36329114 36328027	36330190	32	-	-
NHP2L1	38699406 38699381 38692986 38690899 38690898	38699440	34	-	-
TPST2	23682152 23633783	23682188	36	-	-
IL17R	14510300	14510262	38	+	+
G22P1	38631829 38631856 38631884	38631789	40	+	+
DGCR2	16050147	16050187	40	-	-

	16050106				
SOX10	34995070 34994907 34994406	34995112	42	-	-
RABL2B	47726077 47726068 47726056 47726047 47715995	47726122	45	-	-
FBXO7	29566757 29567034 29571015 29583821	29566711	46	+	+
CRYBB1	23710090	23710139	49	-	-
TEF	38392544 38392655	38392495	49	+	+
GP1BB	16651122 16651293	16651072	50	+	+
PES1	27683923	27683975	52	-	-
SREBF2	38843688	38843633	55	+	+
YWHAH	29036576 29036584	29036518	58	+	+
RANGAP1	38296700 38291578	38296759	59	-	-
H1F0	34815705 34815809	34815645	60	+	+
MFNG	34496916	34496980	64	-	-
MYH9	33427276 33388605 33388586 33325063 33322576 33321670	33427341	65	-	-
CHEK2	25834506 25826811 25826805	25834572	66	-	-
E46L	42703070 42703071 42703080	42703003	67	+	+
EWSR1	26360367 26360379 26360401 26360406 26360415	26360297	70	+	+

	26364301 26382915				
DMC1	35580597 35578823	35580670	73	-	-
NAGA	39081303 39080912	39081376	73	-	-
UBE2L3	18620712 18645843	18620639	73	+	+
TCN2	27699209 27699239 27699240 27699270 27699314	27699131	78	+	+
LGALS1	34686174 34686192	34686079	95	+	+
MTMR3	26975287 27070479	26975188	99	+	+
RANBP1	17045325 17045372	17045221	104	+	+
DRG1	28491652 28491654 28491657	28491547	105	+	+
UFD1L	16406914 16406893 16399457	16407046	132	-	-
CSF2RB	33932753	33932607	146	+	+
SEZ6L	23261735 23384316	23261579	156	+	+
ST13	37867217 37867149 37835780	37867378	161	-	-
BIK	40109209 40122455	40109041	168	+	+
SLC5A1	29135306	29135128	178	+	+
LGALS2	34590469 34590464	34590671	202	-	-
PRKCABP	35067942 35067956 35067974 35067980 35069779 35085930	35067731	211	+	+
NDUFA6	39101279	39101493	214	-	-

SF3A1	27448953 27448925	27449170	217	-	-
FBLN1	42534087 42534180	42533858	229	+	+
RFPL1	26530668	26530432	236	+	+
NUP50	42195347 42195487 42195876 42215805	42195109	238	+	+
MIF	20932709 20932724 20932731 20932741 20932754 20932774 20932807 20932832	20932471	238	+	+
MB	32656679 32655065	32656926	247	-	-
PITPNB	25011314	25011583	269	-	-
EIF3S7	33539727 33539714	33540014	287	-	-
ATF4	36531099 36531116 36531118 36531123 36531889	36530769	330	+	+
RRP22	26407652	26408011	359	-	-
GGA1	34619373 34619384 34619394 34619395	34619013	360	+	+
LZTR1	18035385	18035009	376	+	+
TXN2	33492179 33491415	33492558	379	-	-
NCF4	33871561 33871589 33871686 33871709	33871116	445	+	+
SNRPD3	21648086 21648096	21647555	531	+	+
HMOX1	32473133 32473147	32472400	733	+	+
PPARA	43114051	43113186	865	+	+

	43138914 43138925				
CLDN5	16452151 16452123	16453116	965	-	-
CACNG2	33713152	33714134	982	-	-
BRD1	46807661 46757221	46808757	1096	-	-
SH3BP1	34651664 34651909	34650205	1459	+	+
CRYBB2	22313484	22311707	1777	+	+
SSTR3	34218373	34221536	3163	-	-
CECR1	14634843 14624579	14644379	9536	-	-
TOB2	38448093 38445488	38457814	9721	-	-
ADORA2A	21519741 21525230	21509847	9894	+	+
KCNJ4	35454563 35438706	35465709	11146	-	-
ARHGAP8	41745795 41756450 41762277 41765312 41783762 41817661	41733458	12337	+	+
PIK4CA	17891768 17795662 17787705	17911696	19928	-	-
AP1B1	26459355	26480604	21249	-	-
GNAZ	20136567 20136571	20110695	25872	+	+
SYN3	30098729	30150402	51673	-	-

iv) Genes not annotated in Sanger's current Gene List Table

Gene Symbol	SOE1		Difference	Strand	
	FIE2	Sanger		FIE2	Sanger
MAP3K7IP1	36410325	Not Found		+	
GAS2L1	26399127 26399138 26399141	Not Found		+	
RABL4	33786688 33786645	Not Found		-	
CDC42EP1	34571031	Not Found		+	

	34571062				
ARFGAP3	39855678 39855652 39833842 39830037	Not Found		-	
C22orf19	26645783 26645766 26645757	Not Found		-	

APPENDIX 4

Below are the coordinates of the transcription factor binding sites (TFBSs) listed in Tables 3.1a, b and c within each of the 9 GRIA promoters that were studied. The "+" and "-" signs within brackets indicate the strand orientation of the TFBS.

A.4.1 Coordinates for Top 3 Ranked Singles within the GRIA promoters

Human GRIA1

	Bach2 +
1	224-234(+)

	Sp3 +
1	-107--94(+)
2	376-389(+)

	CDP -
1	93-102(-)
2	218-227(-)
3	484-493(-)

Human GRIA2

	Bach2 +
1	-559--549(+)
2	-11--1(+)

	Sp3 +
1	-325--312(+)
2	-173--160(+)

	CDP -
1	-1148--1139(-)
2	-1136--1127(-)
3	-1136--1127(-)

Human GRIA3

	Bach2 +
1	295-305(+)

```
         Sp3  +
============================
1       -94--81(+)
2       120-133(+)
3       807-820(+)

         CDP  -
============================
1       -1113--1104(-)
2       -1113--1104(-)
3       -561--552(-)
4       -445--436(-)
```

Human GRIA4

```
         Bach2  +
============================
1       272-282(+)

         Sp3  +
============================
1       74-87(+)
2       373-386(+)

         CDP  -
============================
1       -103--94(-)
2       709-718(-)
```


Murine GRIA1

```
         Bach2  +
============================
1       -929--919(+)

         Sp3  +
============================
1       -1110--1097(+)
2       -1109--1096(+)
3       952-965(+)

         CDP  -
============================
1       248-257(-)
2       248-257(-)
```

Murine GRIA2

```
         Bach2  +
============================
1       -1475--1465(+)
2       -435--425(+)

         Sp3  +
```

```
============================
1       -709--696(+)
2       -80--67(+)

         CDP  -
============================
1       -810--801(-)
```

Murine GRIA3

```
         Bach2  +
============================
1       98-108(+)

         Sp3  +
============================
1       -1427--1414(+)
2       -265--252(+)
3       -78--65(+)
4       335-348(+)
5       769-782(+)
6       960-973(+)

         CDP  -
============================
1       -730--721(-)
```


Murine GRIA4

```
         Bach2  +
============================
1       -1310--1300(+)

         Sp3  +
============================
1       -223--210(+)
2       -174--161(+)

         CDP  -
============================
1       -389--380(-)
2       -389--380(-)
3       477-486(-)
```

Rat GRIA1

```
         Bach2  +
============================
1       -1084--1074(+)
2       376-386(+)
3       850-860(+)

         Sp3  +
============================
```

```
1       -726--713(+)
2       -361--348(+)
3       49-62(+)
4       221-234(+)

         CDP  -
==============================
1       245-254(-)
```

A.4.2 Coordinates for Top 3 Ranked Pairs within the GRIA promoters

Human GRIA1

	+ GKLF	+ PU.1
1	-1421--1408(+)	-287--280(+)
2	-841--828(+)	416-423(+)
3	-840--827(+)	
4	-839--826(+)	
5	-838--825(+)	
6	-778--765(+)	
7	-777--764(+)	
8	-604--591(+)	
9	-410--397(+)	
10	-409--396(+)	
11	-406--393(+)	
12	-405--392(+)	
13	-397--384(+)	
14	-388--375(+)	
15	-318--305(+)	
16	-315--302(+)	
17	-309--296(+)	
18	-299--286(+)	
19	-292--279(+)	
20	-291--278(+)	
21	-290--277(+)	
22	-289--276(+)	
23	-283--270(+)	
24	-278--265(+)	
25	-275--262(+)	
26	-184--171(+)	
27	-183--170(+)	
28	-180--167(+)	
29	-179--166(+)	
30	-170--157(+)	
31	-166--153(+)	
32	-151--138(+)	
33	-146--133(+)	
34	-145--132(+)	
35	-140--127(+)	
36	-139--126(+)	
37	-138--125(+)	
38	-137--124(+)	
39	-107--94(+)	
40	-4-9(+)	
41	12-25(+)	
42	16-29(+)	
43	39-52(+)	
44	62-75(+)	
45	63-76(+)	
46	68-81(+)	
47	69-82(+)	
48	70-83(+)	
49	71-84(+)	

50	72-85(+)
51	73-86(+)
52	126-139(+)
53	224-237(+)
54	348-361(+)
55	357-370(+)
56	374-387(+)
57	375-388(+)
58	376-389(+)
59	377-390(+)
60	399-412(+)
61	407-420(+)
62	501-514(+)
63	510-523(+)
64	511-524(+)
65	706-719(+)
66	707-720(+)
67	708-721(+)

Pairs:

15-1	-318--305(+)	-287--280(+)
16-1	-315--302(+)	-287--280(+)
17-1	-309--296(+)	-287--280(+)
55-2	357-370(+)	416-423(+)
56-2	374-387(+)	416-423(+)
57-2	375-388(+)	416-423(+)
58-2	376-389(+)	416-423(+)
59-2	377-390(+)	416-423(+)
60-2	399-412(+)	416-423(+)

****

	+ MZF1	+ GATA-2
1	-501--494(+)	-1483--1474(+)
2	-433--426(+)	-1166--1157(+)
3	-402--395(+)	-577--568(+)
4	-400--388(+)	-305--296(+)
5	-287--275(+)	281-290(+)
6	-286--274(+)	
7	-282--275(+)	
8	-200--188(+)	
9	-196--189(+)	
10	-174--162(+)	
11	-152--140(+)	
12	-149--142(+)	
13	-130--123(+)	
14	76-88(+)	
15	77-89(+)	
16	80-87(+)	
17	227-239(+)	
18	241-248(+)	
19	278-285(+)	
20	282-294(+)	
21	380-392(+)	
22	403-415(+)	
23	411-423(+)	
24	414-421(+)	

25 679-686(+)
26 712-724(+)

Pairs:

17-5 227-239(+) 281-290(+)
18-5 241-248(+) 281-290(+)

****

	+ PU.1	+ GKLF
1	-287--280(+)	-1421--1408(+)
2	416-423(+)	-841--828(+)
3		-840--827(+)
4		-839--826(+)
5		-838--825(+)
6		-778--765(+)
7		-777--764(+)
8		-604--591(+)
9		-410--397(+)
10		-409--396(+)
11		-406--393(+)
12		-405--392(+)
13		-397--384(+)
14		-388--375(+)
15		-318--305(+)
16		-315--302(+)
17		-309--296(+)
18		-299--286(+)
19		-292--279(+)
20		-291--278(+)
21		-290--277(+)
22		-289--276(+)
23		-283--270(+)
24		-278--265(+)
25		-275--262(+)
26		-184--171(+)
27		-183--170(+)
28		-180--167(+)
29		-179--166(+)
30		-170--157(+)
31		-166--153(+)
32		-151--138(+)
33		-146--133(+)
34		-145--132(+)
35		-140--127(+)
36		-139--126(+)
37		-138--125(+)
38		-137--124(+)
39		-107--94(+)
40		-4-9(+)
41		12-25(+)
42		16-29(+)
43		39-52(+)
44		62-75(+)
45		63-76(+)
46		68-81(+)
47		69-82(+)

48	70-83(+)
49	71-84(+)
50	72-85(+)
51	73-86(+)
52	126-139(+)
53	224-237(+)
54	348-361(+)
55	357-370(+)
56	374-387(+)
57	375-388(+)
58	376-389(+)
59	377-390(+)
60	399-412(+)
61	407-420(+)
62	501-514(+)
63	510-523(+)
64	511-524(+)
65	706-719(+)
66	707-720(+)
67	708-721(+)

Pairs:

1-24	-287--280(+)	-278--265(+)
1-25	-287--280(+)	-275--262(+)

****

Human GRIA2

	+ GKLF	+ PU.1
1	-1406--1393(+)	269-276(+)
2	-1012--999(+)	
3	-1011--998(+)	
4	-870--857(+)	
5	-830--817(+)	
6	-741--728(+)	
7	-545--532(+)	
8	-468--455(+)	
9	-148--135(+)	
10	-147--134(+)	
11	-3-10(+)	
12	-2-11(+)	
13	18-31(+)	
14	19-32(+)	
15	45-58(+)	
16	136-149(+)	
17	137-150(+)	
18	178-191(+)	
19	179-192(+)	
20	244-257(+)	
21	277-290(+)	
22	283-296(+)	
23	286-299(+)	
24	292-305(+)	
25	303-316(+)	
26	304-317(+)	
27	305-318(+)	

28	590-603(+)
29	625-638(+)
30	735-748(+)

Pairs:

20-1 244-257(+) 269-276(+)

****

	+ MZF1	+ GATA-2
1	-1473--1466(+)	-1434--1425(+)
2	-1008--996(+)	-1389--1380(+)
3	-1004--997(+)	-1007--998(+)
4	-318--311(+)	-496--487(+)
5	-146--139(+)	-496--487(+)
6	-144--132(+)	429-438(+)
7	3-15(+)	
8	24-36(+)	
9	138-145(+)	
10	289-301(+)	
11	309-321(+)	
12	388-400(+)	
13	392-399(+)	
14	982-994(+)	

Pairs:

1-1	-1473--1466(+)	-1434--1425(+)
12-6	388-400(+)	429-438(+)
13-6	392-399(+)	429-438(+)

****

	+ PU.1	+ GKLF
1	269-276(+)	-1406--1393(+)
2		-1012--999(+)
3		-1011--998(+)
4		-870--857(+)
5		-830--817(+)
6		-741--728(+)
7		-545--532(+)
8		-468--455(+)
9		-148--135(+)
10		-147--134(+)
11		-3-10(+)
12		-2-11(+)
13		18-31(+)
14		19-32(+)
15		45-58(+)
16		136-149(+)
17		137-150(+)
18		178-191(+)
19		179-192(+)
20		244-257(+)
21		277-290(+)
22		283-296(+)
23		286-299(+)

24	292-305(+)
25	303-316(+)
26	304-317(+)
27	305-318(+)
28	590-603(+)
29	625-638(+)
30	735-748(+)

Pairs:

1-21	269-276(+)	277-290(+)
1-22	269-276(+)	283-296(+)
1-23	269-276(+)	286-299(+)
1-24	269-276(+)	292-305(+)
1-25	269-276(+)	303-316(+)
1-26	269-276(+)	304-317(+)
1-27	269-276(+)	305-318(+)

****

Human GRIA3

	+ GKLF	+ PU.1
1	-1457--1444(+)	-1197--1190(+)
2	-1446--1433(+)	-82--75(+)
3	-1362--1349(+)	42-49(+)
4	-1333--1320(+)	
5	-1332--1319(+)	
6	-1327--1314(+)	
7	-1326--1313(+)	
8	-1206--1193(+)	
9	-870--857(+)	
10	-869--856(+)	
11	-834--821(+)	
12	-791--778(+)	
13	-652--639(+)	
14	-622--609(+)	
15	-339--326(+)	
16	-338--325(+)	
17	-337--324(+)	
18	-98--85(+)	
19	-94--81(+)	
20	-93--80(+)	
21	-92--79(+)	
22	-87--74(+)	
23	-86--73(+)	
24	-83--70(+)	
25	-82--69(+)	
26	-81--68(+)	
27	-80--67(+)	
28	-54--41(+)	
29	-45--32(+)	
30	-44--31(+)	
31	-33--20(+)	
32	17-30(+)	
33	18-31(+)	
34	33-46(+)	
35	42-55(+)	

36	45-58(+)
37	48-61(+)
38	66-79(+)
39	69-82(+)
40	89-102(+)
41	114-127(+)
42	115-128(+)
43	116-129(+)
44	119-132(+)
45	120-133(+)
46	121-134(+)
47	122-135(+)
48	238-251(+)
49	239-252(+)
50	240-253(+)
51	443-456(+)
52	482-495(+)
53	542-555(+)
54	543-556(+)
55	544-557(+)
56	620-633(+)
57	632-645(+)
58	882-895(+)
59	883-896(+)
60	884-897(+)

Pairs:

18-2	-98--85(+)	-82--75(+)
32-3	17-30(+)	42-49(+)
33-3	18-31(+)	42-49(+)

****

	+ MZF1	+ GATA-2
1	-1421--1414(+)	-1447--1438(+)
2	-1328--1316(+)	-1361--1352(+)
3	-1202--1190(+)	-836--827(+)
4	-1199--1192(+)	724-733(+)
5	-866--859(+)	
6	-864--852(+)	
7	-839--832(+)	
8	-743--731(+)	
9	-88--76(+)	
10	-87--75(+)	
11	-84--77(+)	
12	-77--65(+)	
13	-49--37(+)	
14	23-35(+)	
15	123-130(+)	
16	125-137(+)	
17	493-505(+)	
18	496-503(+)	
19	548-560(+)	
20	816-823(+)	
21	874-886(+)	

Pairs:

5-3	-866--859(+)	-836--827(+)
6-3	-864--852(+)	-836--827(+)

****

	+ PU.1	+ GKLF
1	-1197--1190(+)	-1457--1444(+)
2	-82--75(+)	-1446--1433(+)
3	42-49(+)	-1362--1349(+)
4		-1333--1320(+)
5		-1332--1319(+)
6		-1327--1314(+)
7		-1326--1313(+)
8		-1206--1193(+)
9		-870--857(+)
10		-869--856(+)
11		-834--821(+)
12		-791--778(+)
13		-652--639(+)
14		-622--609(+)
15		-339--326(+)
16		-338--325(+)
17		-337--324(+)
18		-98--85(+)
19		-94--81(+)
20		-93--80(+)
21		-92--79(+)
22		-87--74(+)
23		-86--73(+)
24		-83--70(+)
25		-82--69(+)
26		-81--68(+)
27		-80--67(+)
28		-54--41(+)
29		-45--32(+)
30		-44--31(+)
31		-33--20(+)
32		17-30(+)
33		18-31(+)
34		33-46(+)
35		42-55(+)
36		45-58(+)
37		48-61(+)
38		66-79(+)
39		69-82(+)
40		89-102(+)
41		114-127(+)
42		115-128(+)
43		116-129(+)
44		119-132(+)
45		120-133(+)
46		121-134(+)
47		122-135(+)
48		238-251(+)
49		239-252(+)
50		240-253(+)
51		443-456(+)

```
52                                  482-495(+)
53                                  542-555(+)
54                                  543-556(+)
55                                  544-557(+)
56                                  620-633(+)
57                                  632-645(+)
58                                  882-895(+)
59                                  883-896(+)
60                                  884-897(+)

Pairs:
2-28    -82--75(+)                  -54--41(+)
2-29    -82--75(+)                  -45--32(+)
2-30    -82--75(+)                  -44--31(+)
2-31    -82--75(+)                  -33--20(+)
3-38    42-49(+)                    66-79(+)
3-39    42-49(+)                    69-82(+)
3-40    42-49(+)                    89-102(+)
************************************************
```

Human GRIA4

```
            + GKLF                  + PU.1
================================================
1       -1429--1416(+)              -865--858(+)
2       -874--861(+)                256-263(+)
3       -860--847(+)
4       -789--776(+)
5       -6-7(+)
6       -5-8(+)
7       0-13(+)
8       9-22(+)
9       17-30(+)
10      30-43(+)
11      31-44(+)
12      34-47(+)
13      75-88(+)
14      199-212(+)
15      200-213(+)
16      201-214(+)
17      258-271(+)
18      261-274(+)
19      262-275(+)
20      284-297(+)
21      411-424(+)
22      423-436(+)
23      535-548(+)
24      536-549(+)
25      537-550(+)
26      911-924(+)
```

Pairs:

```
14-2    199-212(+)                  256-263(+)
15-2    200-213(+)                  256-263(+)
16-2    201-214(+)                  256-263(+)
************************************************
```

	+ MZF1	+ GATA-2
1	97-104(+)	-721--712(+)
2	205-217(+)	469-478(+)
3	288-300(+)	573-582(+)
4	384-396(+)	
5	415-427(+)	
6	418-425(+)	
7	466-473(+)	
8	541-553(+)	
9	901-913(+)	
10	962-969(+)	

Pairs:

5-2	415-427(+)	469-478(+)
6-2	418-425(+)	469-478(+)
8-3	541-553(+)	573-582(+)

****

	+ PU.1	+ GKLF
1	-865--858(+)	-1429--1416(+)
2	256-263(+)	-874--861(+)
3		-860--847(+)
4		-789--776(+)
5		-6-7(+)
6		-5-8(+)
7		0-13(+)
8		9-22(+)
9		17-30(+)
10		30-43(+)
11		31-44(+)
12		34-47(+)
13		75-88(+)
14		199-212(+)
15		200-213(+)
16		201-214(+)
17		258-271(+)
18		261-274(+)
19		262-275(+)
20		284-297(+)
21		411-424(+)
22		423-436(+)
23		535-548(+)
24		536-549(+)
25		537-550(+)
26		911-924(+)

Pairs:

2-20	256-263(+)	284-297(+)

****

Murine GRIA1

	+ GKLF	+ PU.1

1	-1478--1465(+)	-1091--1084(+)
2	-1387--1374(+)	-947--940(+)
3	-1317--1304(+)	-628--621(+)
4	-1289--1276(+)	-12--5(+)
5	-1268--1255(+)	
6	-1267--1254(+)	
7	-1266--1253(+)	
8	-1265--1252(+)	
9	-1264--1251(+)	
10	-1259--1246(+)	
11	-1258--1245(+)	
12	-1255--1242(+)	
13	-1254--1241(+)	
14	-1250--1237(+)	
15	-1246--1233(+)	
16	-1234--1221(+)	
17	-1230--1217(+)	
18	-1226--1213(+)	
19	-1222--1209(+)	
20	-1218--1205(+)	
21	-1214--1201(+)	
22	-1210--1197(+)	
23	-1206--1193(+)	
24	-1202--1189(+)	
25	-1197--1184(+)	
26	-1194--1181(+)	
27	-1181--1168(+)	
28	-1178--1165(+)	
29	-1177--1164(+)	
30	-1116--1103(+)	
31	-1115--1102(+)	
32	-1114--1101(+)	
33	-1111--1098(+)	
34	-1110--1097(+)	
35	-1109--1096(+)	
36	-1108--1095(+)	
37	-1107--1094(+)	
38	-1104--1091(+)	
39	-1103--1090(+)	
40	-1100--1087(+)	
41	-1032--1019(+)	
42	-1028--1015(+)	
43	-1024--1011(+)	
44	-1012--999(+)	
45	-968--955(+)	
46	-967--954(+)	
47	-966--953(+)	
48	-963--950(+)	
49	-960--947(+)	
50	-959--946(+)	
51	-956--943(+)	
52	-952--939(+)	
53	-951--938(+)	
54	-749--736(+)	
55	-658--645(+)	
56	-657--644(+)	
57	-656--643(+)	

58	-637--624(+)
59	-472--459(+)
60	-455--442(+)
61	-454--441(+)
62	-196--183(+)
63	-192--179(+)
64	-191--178(+)
65	-190--177(+)
66	-145--132(+)
67	-74--61(+)
68	-60--47(+)
69	32-45(+)
70	344-357(+)
71	611-624(+)
72	617-630(+)
73	843-856(+)
74	905-918(+)

Pairs:

30-1	-1116--1103(+)	-1091--1084(+)
31-1	-1115--1102(+)	-1091--1084(+)
32-1	-1114--1101(+)	-1091--1084(+)
33-1	-1111--1098(+)	-1091--1084(+)
34-1	-1110--1097(+)	-1091--1084(+)
35-1	-1109--1096(+)	-1091--1084(+)
36-1	-1108--1095(+)	-1091--1084(+)
37-1	-1107--1094(+)	-1091--1084(+)
45-2	-968--955(+)	-947--940(+)
46-2	-967--954(+)	-947--940(+)
47-2	-966--953(+)	-947--940(+)
48-2	-963--950(+)	-947--940(+)
55-3	-658--645(+)	-628--621(+)
56-3	-657--644(+)	-628--621(+)
57-3	-656--643(+)	-628--621(+)
67-4	-74--61(+)	-12--5(+)
68-4	-60--47(+)	-12--5(+)

****

	+ MZF1	+ GATA-2
1	-1463--1451(+)	-646--637(+)
2	-1262--1250(+)	-600--591(+)
3	-1257--1250(+)	-595--586(+)
4	-1107--1100(+)	-544--535(+)
5	-1105--1093(+)	
6	-1100--1093(+)	
7	-1098--1086(+)	
8	-959--952(+)	
9	-947--935(+)	
10	-652--640(+)	
11	-649--642(+)	
12	-187--175(+)	
13	-183--176(+)	
14	-68--56(+)	
15	152-164(+)	
16	156-163(+)	
17	616-628(+)	

18 908-920(+)
19 912-919(+)

Pairs:

10-2	-652--640(+)	-600--591(+)
10-3	-652--640(+)	-595--586(+)
11-2	-649--642(+)	-600--591(+)
11-3	-649--642(+)	-595--586(+)

****

	+ PU.1	+ GKLF
1	-1091--1084(+)	-1478--1465(+)
2	-947--940(+)	-1387--1374(+)
3	-628--621(+)	-1317--1304(+)
4	-12--5(+)	-1289--1276(+)
5		-1268--1255(+)
6		-1267--1254(+)
7		-1266--1253(+)
8		-1265--1252(+)
9		-1264--1251(+)
10		-1259--1246(+)
11		-1258--1245(+)
12		-1255--1242(+)
13		-1254--1241(+)
14		-1250--1237(+)
15		-1246--1233(+)
16		-1234--1221(+)
17		-1230--1217(+)
18		-1226--1213(+)
19		-1222--1209(+)
20		-1218--1205(+)
21		-1214--1201(+)
22		-1210--1197(+)
23		-1206--1193(+)
24		-1202--1189(+)
25		-1197--1184(+)
26		-1194--1181(+)
27		-1181--1168(+)
28		-1178--1165(+)
29		-1177--1164(+)
30		-1116--1103(+)
31		-1115--1102(+)
32		-1114--1101(+)
33		-1111--1098(+)
34		-1110--1097(+)
35		-1109--1096(+)
36		-1108--1095(+)
37		-1107--1094(+)
38		-1104--1091(+)
39		-1103--1090(+)
40		-1100--1087(+)
41		-1032--1019(+)
42		-1028--1015(+)
43		-1024--1011(+)
44		-1012--999(+)
45		-968--955(+)

46		-967--954(+)
47		-966--953(+)
48		-963--950(+)
49		-960--947(+)
50		-959--946(+)
51		-956--943(+)
52		-952--939(+)
53		-951--938(+)
54		-749--736(+)
55		-658--645(+)
56		-657--644(+)
57		-656--643(+)
58		-637--624(+)
59		-472--459(+)
60		-455--442(+)
61		-454--441(+)
62		-196--183(+)
63		-192--179(+)
64		-191--178(+)
65		-190--177(+)
66		-145--132(+)
67		-74--61(+)
68		-60--47(+)
69		32-45(+)
70		344-357(+)
71		611-624(+)
72		617-630(+)
73		843-856(+)
74		905-918(+)

Pairs:

4-69	-12--5(+)	32-45(+)

****

Murine GRIA2

	+ GKLF	+ PU.1
1	-1500--1487(+)	359-366(+)
2	-1497--1484(+)	
3	-1472--1459(+)	
4	-1471--1458(+)	
5	-1468--1455(+)	
6	-1339--1326(+)	
7	-1050--1037(+)	
8	-848--835(+)	
9	-706--693(+)	
10	-355--342(+)	
11	-354--341(+)	
12	-207--194(+)	
13	-55--42(+)	
14	-51--38(+)	
15	91-104(+)	
16	92-105(+)	
17	113-126(+)	
18	137-150(+)	

```
19      221-234(+)
20      226-239(+)
21      334-347(+)
22      373-386(+)
23      376-389(+)
24      381-394(+)
25      382-395(+)
26      390-403(+)
27      391-404(+)
28      392-405(+)
29      709-722(+)
30      710-723(+)
31      715-728(+)
32      820-833(+)
```

Pairs:

```
21-1    334-347(+)              359-366(+)
**************************************************
```

```
            + MZF1              + GATA-2
==================================================
1       -1493--1486(+)          -1270--1261(+)
2       -1491--1479(+)          -1044--1035(+)
3       -1468--1456(+)          -925--916(+)
4       -699--692(+)            -815--806(+)
5       -631--619(+)            -540--531(+)
6       -53--46(+)              152-161(+)
7       -51--39(+)              272-281(+)
8       97-109(+)               272-281(+)
9       118-130(+)              402-411(+)
10      396-408(+)              517-526(+)
11      399-406(+)
12      476-488(+)
13      480-487(+)
14      714-726(+)
15      717-724(+)
```


Pairs:

```
8-6      97-109(+)              152-161(+)
9-6      118-130(+)             152-161(+)
12-10    476-488(+)             517-526(+)
13-10    480-487(+)             517-526(+)
**************************************************
```

```
            + PU.1              + GKLF
==================================================
1       359-366(+)              -1500--1487(+)
2                               -1497--1484(+)
3                               -1472--1459(+)
4                               -1471--1458(+)
5                               -1468--1455(+)
6                               -1339--1326(+)
7                               -1050--1037(+)
8                               -848--835(+)
9                               -706--693(+)
10                              -355--342(+)
11                              -354--341(+)
```

12		-207--194(+)
13		-55--42(+)
14		-51--38(+)
15		91-104(+)
16		92-105(+)
17		113-126(+)
18		137-150(+)
19		221-234(+)
20		226-239(+)
21		334-347(+)
22		373-386(+)
23		376-389(+)
24		381-394(+)
25		382-395(+)
26		390-403(+)
27		391-404(+)
28		392-405(+)
29		709-722(+)
30		710-723(+)
31		715-728(+)
32		820-833(+)

Pairs:

1-22	359-366(+)	373-386(+)
1-23	359-366(+)	376-389(+)
1-24	359-366(+)	381-394(+)
1-25	359-366(+)	382-395(+)
1-26	359-366(+)	390-403(+)
1-27	359-366(+)	391-404(+)
1-28	359-366(+)	392-405(+)

****

Murine GRIA3

	+ GKLF	+ PU.1
1	-1484--1471(+)	-211--204(+)
2	-1480--1467(+)	-196--189(+)
3	-1471--1458(+)	-171--164(+)
4	-1427--1414(+)	885-892(+)
5	-1379--1366(+)	
6	-1378--1365(+)	
7	-1377--1364(+)	
8	-1166--1153(+)	
9	-1123--1110(+)	
10	-901--888(+)	
11	-269--256(+)	
12	-265--252(+)	
13	-264--251(+)	
14	-263--250(+)	
15	-251--238(+)	
16	-228--215(+)	
17	-227--214(+)	
18	-220--207(+)	
19	-212--199(+)	
20	-205--192(+)	

21 -197--184(+)
22 -192--179(+)
23 -191--178(+)
24 -180--167(+)
25 -154--141(+)
26 -151--138(+)
27 -148--135(+)
28 -145--132(+)
29 -113--100(+)
30 -82--69(+)
31 -79--66(+)
32 -78--65(+)
33 -77--64(+)
34 41-54(+)
35 42-55(+)
36 43-56(+)
37 156-169(+)
38 157-170(+)
39 260-273(+)
40 330-343(+)
41 334-347(+)
42 335-348(+)
43 336-349(+)
44 389-402(+)
45 390-403(+)
46 409-422(+)
47 603-616(+)
48 657-670(+)
49 658-671(+)
50 662-675(+)
51 670-683(+)
52 688-701(+)
53 769-782(+)
54 865-878(+)
55 875-888(+)
56 955-968(+)

Pairs:

11-1	-269--256(+)	-211--204(+)
12-1	-265--252(+)	-211--204(+)
13-1	-264--251(+)	-211--204(+)
14-1	-263--250(+)	-211--204(+)
15-1	-251--238(+)	-211--204(+)
15-2	-251--238(+)	-196--189(+)
16-1	-228--215(+)	-211--204(+)
16-2	-228--215(+)	-196--189(+)
16-3	-228--215(+)	-171--164(+)
17-1	-227--214(+)	-211--204(+)
17-2	-227--214(+)	-196--189(+)
17-3	-227--214(+)	-171--164(+)
18-2	-220--207(+)	-196--189(+)
18-3	-220--207(+)	-171--164(+)
19-2	-212--199(+)	-196--189(+)
19-3	-212--199(+)	-171--164(+)
20-3	-205--192(+)	-171--164(+)
21-3	-197--184(+)	-171--164(+)
22-3	-192--179(+)	-171--164(+)

```
23-3     -191--178(+)             -171--164(+)
54-4     865-878(+)               885-892(+)
************************************************

            + MZF1                + GATA-2
================================================
1       -1424--1417(+)            -1133--1124(+)
2       -1422--1410(+)            264-273(+)
3       -1373--1361(+)            282-291(+)
4       -1167--1155(+)            512-521(+)
5       -1163--1156(+)            662-671(+)
6       -1161--1149(+)            683-692(+)
7       -1136--1129(+)
8       -1134--1122(+)
9       -1129--1122(+)
10      -1023--1011(+)
11      -619--607(+)
12      -615--608(+)
13      -320--313(+)
14      -247--235(+)
15      -223--211(+)
16      -218--211(+)
17      -186--174(+)
18      -78--66(+)
19      -75--68(+)
20      -73--61(+)
21      160-172(+)
22      248-255(+)
23      263-275(+)
24      268-275(+)
25      316-328(+)
26      320-327(+)
27      388-400(+)
28      661-673(+)
29      860-872(+)
30      863-870(+)
31      880-892(+)
32      883-890(+)
33      928-940(+)
```


Pairs:

```
4-1      -1167--1155(+)           -1133--1124(+)
5-1      -1163--1156(+)           -1133--1124(+)
6-1      -1161--1149(+)           -1133--1124(+)
22-2     248-255(+)               264-273(+)
22-3     248-255(+)               282-291(+)
23-3     263-275(+)               282-291(+)
24-3     268-275(+)               282-291(+)
28-6     661-673(+)               683-692(+)
************************************************

            + PU.1                + GKLF
================================================
1       -211--204(+)              -1484--1471(+)
2       -196--189(+)              -1480--1467(+)
3       -171--164(+)              -1471--1458(+)
4       885-892(+)                -1427--1414(+)
```

5	-1379--1366(+)
6	-1378--1365(+)
7	-1377--1364(+)
8	-1166--1153(+)
9	-1123--1110(+)
10	-901--888(+)
11	-269--256(+)
12	-265--252(+)
13	-264--251(+)
14	-263--250(+)
15	-251--238(+)
16	-228--215(+)
17	-227--214(+)
18	-220--207(+)
19	-212--199(+)
20	-205--192(+)
21	-197--184(+)
22	-192--179(+)
23	-191--178(+)
24	-180--167(+)
25	-154--141(+)
26	-151--138(+)
27	-148--135(+)
28	-145--132(+)
29	-113--100(+)
30	-82--69(+)
31	-79--66(+)
32	-78--65(+)
33	-77--64(+)
34	41-54(+)
35	42-55(+)
36	43-56(+)
37	156-169(+)
38	157-170(+)
39	260-273(+)
40	330-343(+)
41	334-347(+)
42	335-348(+)
43	336-349(+)
44	389-402(+)
45	390-403(+)
46	409-422(+)
47	603-616(+)
48	657-670(+)
49	658-671(+)
50	662-675(+)
51	670-683(+)
52	688-701(+)
53	769-782(+)
54	865-878(+)
55	875-888(+)
56	955-968(+)

Pairs:

1-21	-211--204(+)	-197--184(+)
1-22	-211--204(+)	-192--179(+)
1-23	-211--204(+)	-191--178(+)

1-24	-211--204(+)	-180--167(+)
1-25	-211--204(+)	-154--141(+)
2-24	-196--189(+)	-180--167(+)
2-25	-196--189(+)	-154--141(+)
2-26	-196--189(+)	-151--138(+)
2-27	-196--189(+)	-148--135(+)
2-28	-196--189(+)	-145--132(+)
3-25	-171--164(+)	-154--141(+)
3-26	-171--164(+)	-151--138(+)
3-27	-171--164(+)	-148--135(+)
3-28	-171--164(+)	-145--132(+)
3-29	-171--164(+)	-113--100(+)

Murine GRIA4

	+ GKLF	+ PU.1
1	-751--738(+)	-1264--1257(+)
2	-229--216(+)	-225--218(+)
3	-228--215(+)	34-41(+)
4	-223--210(+)	45-52(+)
5	-155--142(+)	804-811(+)
6	-58--45(+)	979-986(+)
7	-26--13(+)	
8	40-53(+)	
9	41-54(+)	
10	42-55(+)	
11	137-150(+)	
12	144-157(+)	
13	145-158(+)	
14	150-163(+)	
15	151-164(+)	
16	154-167(+)	
17	159-172(+)	
18	166-179(+)	
19	167-180(+)	
20	168-181(+)	
21	173-186(+)	
22	174-187(+)	
23	175-188(+)	
24	181-194(+)	
25	182-195(+)	
26	310-323(+)	
27	311-324(+)	
28	320-333(+)	
29	321-334(+)	
30	660-673(+)	
31	681-694(+)	
32	697-710(+)	
33	793-806(+)	
34	794-807(+)	
35	795-808(+)	
36	801-814(+)	
37	970-983(+)	
38	974-987(+)	

Pairs:

7-3	-26--13(+)	34-41(+)

****

	+ MZF1	+ GATA-2
1	-166--159(+)	-1409--1400(+)
2	-151--144(+)	-1409--1400(+)
3	-9-3(+)	-1110--1101(+)
4	45-57(+)	57-66(+)
5	49-56(+)	845-854(+)
6	143-155(+)	845-854(+)
7	146-153(+)	
8	165-177(+)	
9	172-184(+)	
10	179-191(+)	
11	186-198(+)	
12	473-480(+)	
13	486-498(+)	
14	687-699(+)	
15	748-755(+)	
16	802-809(+)	

Pairs:

5-4	49-56(+)	57-66(+)
16-5	802-809(+)	845-854(+)
16-6	802-809(+)	845-854(+)

****

	+ PU.1	+ GKLF
1	-1264--1257(+)	-751--738(+)
2	-225--218(+)	-229--216(+)
3	34-41(+)	-228--215(+)
4	45-52(+)	-223--210(+)
5	804-811(+)	-155--142(+)
6	979-986(+)	-58--45(+)
7		-26--13(+)
8		40-53(+)
9		41-54(+)
10		42-55(+)
11		137-150(+)
12		144-157(+)
13		145-158(+)
14		150-163(+)
15		151-164(+)
16		154-167(+)
17		159-172(+)
18		166-179(+)
19		167-180(+)
20		168-181(+)
21		173-186(+)
22		174-187(+)
23		175-188(+)
24		181-194(+)
25		182-195(+)

26		310-323(+)
27		311-324(+)
28		320-333(+)
29		321-334(+)
30		660-673(+)
31		681-694(+)
32		697-710(+)
33		793-806(+)
34		794-807(+)
35		795-808(+)
36		801-814(+)
37		970-983(+)
38		974-987(+)

Pairs:

3-10	34-41(+)	42-55(+)

****

Rat GRIA1

	+ GKLF	+ PU.1
1	-1377--1364(+)	-501--494(+)
2	-1376--1363(+)	-327--320(+)
3	-737--724(+)	158-165(+)
4	-736--723(+)	524-531(+)
5	-733--720(+)	591-598(+)
6	-732--719(+)	
7	-731--718(+)	
8	-730--717(+)	
9	-729--716(+)	
10	-728--715(+)	
11	-727--714(+)	
12	-726--713(+)	
13	-725--712(+)	
14	-724--711(+)	
15	-723--710(+)	
16	-662--649(+)	
17	-634--621(+)	
18	-524--511(+)	
19	-355--342(+)	
20	-354--341(+)	
21	-349--336(+)	
22	-336--323(+)	
23	-332--319(+)	
24	-328--315(+)	
25	-327--314(+)	
26	-326--313(+)	
27	-223--210(+)	
28	-218--205(+)	
29	-209--196(+)	
30	-205--192(+)	
31	-204--191(+)	
32	-203--190(+)	
33	-114--101(+)	
34	-44--31(+)	

35 -43--30(+)
36 -39--26(+)
37 -30--17(+)
38 -23--10(+)
39 -4-9(+)
40 -1-12(+)
41 9-22(+)
42 10-23(+)
43 15-28(+)
44 16-29(+)
45 17-30(+)
46 18-31(+)
47 49-62(+)
48 145-158(+)
49 146-159(+)
50 149-162(+)
51 153-166(+)
52 169-182(+)
53 192-205(+)
54 217-230(+)
55 220-233(+)
56 221-234(+)
57 222-235(+)
58 225-238(+)
59 226-239(+)
60 278-291(+)
61 376-389(+)
62 377-390(+)
63 390-403(+)
64 512-525(+)
65 525-538(+)
66 544-557(+)
67 545-558(+)
68 546-559(+)
69 573-586(+)
70 582-595(+)
71 795-808(+)
72 879-892(+)

Pairs:

18-1	-524--511(+)	-501--494(+)
19-2	-355--342(+)	-327--320(+)
20-2	-354--341(+)	-327--320(+)
21-2	-349--336(+)	-327--320(+)
66-5	544-557(+)	591-598(+)
67-5	545-558(+)	591-598(+)
68-5	546-559(+)	591-598(+)
69-5	573-586(+)	591-598(+)

****

	+ MZF1	+ GATA-2
1	-727--715(+)	-1314--1305(+)
2	-721--709(+)	-1276--1267(+)
3	-720--708(+)	-661--652(+)
4	-716--709(+)	369-378(+)
5	-421--414(+)	860-869(+)

6	-356--344(+)
7	-352--345(+)
8	-350--338(+)
9	-319--312(+)
10	-312--300(+)
11	-308--301(+)
12	-284--277(+)
13	-232--220(+)
14	-228--221(+)
15	-199--187(+)
16	-183--176(+)
17	-109--97(+)
18	-34--22(+)
19	21-33(+)
20	25-32(+)
21	226-238(+)
22	229-236(+)
23	381-393(+)
24	393-400(+)
25	406-413(+)
26	455-467(+)
27	458-465(+)
28	578-590(+)
29	586-598(+)
30	589-596(+)
31	696-708(+)
32	853-865(+)
33	857-864(+)
34	881-888(+)
35	883-895(+)
36	886-893(+)

Pairs:

2-3	-721--709(+)	-661--652(+)
3-3	-720--708(+)	-661--652(+)
4-3	-716--709(+)	-661--652(+)

****

	+ PU.1	+ GKLF
1	-501--494(+)	-1377--1364(+)
2	-327--320(+)	-1376--1363(+)
3	158-165(+)	-737--724(+)
4	524-531(+)	-736--723(+)
5	591-598(+)	-733--720(+)
6		-732--719(+)
7		-731--718(+)
8		-730--717(+)
9		-729--716(+)
10		-728--715(+)
11		-727--714(+)
12		-726--713(+)
13		-725--712(+)
14		-724--711(+)
15		-723--710(+)
16		-662--649(+)
17		-634--621(+)

18	-524--511(+)
19	-355--342(+)
20	-354--341(+)
21	-349--336(+)
22	-336--323(+)
23	-332--319(+)
24	-328--315(+)
25	-327--314(+)
26	-326--313(+)
27	-223--210(+)
28	-218--205(+)
29	-209--196(+)
30	-205--192(+)
31	-204--191(+)
32	-203--190(+)
33	-114--101(+)
34	-44--31(+)
35	-43--30(+)
36	-39--26(+)
37	-30--17(+)
38	-23--10(+)
39	-4-9(+)
40	-1-12(+)
41	9-22(+)
42	10-23(+)
43	15-28(+)
44	16-29(+)
45	17-30(+)
46	18-31(+)
47	49-62(+)
48	145-158(+)
49	146-159(+)
50	149-162(+)
51	153-166(+)
52	169-182(+)
53	192-205(+)
54	217-230(+)
55	220-233(+)
56	221-234(+)
57	222-235(+)
58	225-238(+)
59	226-239(+)
60	278-291(+)
61	376-389(+)
62	377-390(+)
63	390-403(+)
64	512-525(+)
65	525-538(+)
66	544-557(+)
67	545-558(+)
68	546-559(+)
69	573-586(+)
70	582-595(+)
71	795-808(+)
72	879-892(+)

Pairs:

```
3-52  158-165(+)        169-182(+)
3-53  158-165(+)        192-205(+)
4-66  524-531(+)        544-557(+)
4-67  524-531(+)        545-558(+)
4-68  524-531(+)        546-559(+)
4-69  524-531(+)        573-586(+)
4-70  524-531(+)        582-595(+)
************************************************
```

A.4.3 Coordinates for Top 3 Ranked Triplets within the GRIA promoters

Human GRIA1

```
           - STAT6            + MZF1             - STAT3
============================================================
1      -1462--1455(-)         -501--494(+)       -1462--1455(-)
2      -842--835(-)           -433--426(+)       -851--844(-)
3      -285--278(-)           -402--395(+)       -842--835(-)
4      -20--13(-)             -400--388(+)       -616--609(-)
5      3-10(-)                -287--275(+)       -603--596(-)
6      7-14(-)                -286--274(+)       -285--278(-)
7      23-30(-)               -282--275(+)       -278--271(-)
8      39-46(-)               -200--188(+)       -254--247(-)
9      74-81(-)               -196--189(+)       -20--13(-)
10     84-91(-)               -174--162(+)       3-10(-)
11     418-425(-)             -152--140(+)       7-14(-)
12     497-504(-)             -149--142(+)       23-30(-)
13                            -130--123(+)       27-34(-)
14                            76-88(+)           74-81(-)
15                            77-89(+)           84-91(-)
16                            80-87(+)           137-144(-)
17                            227-239(+)         418-425(-)
18                            241-248(+)         480-487(-)
19                            278-285(+)         867-874(-)
20                            282-294(+)
21                            380-392(+)
22                            403-415(+)
23                            411-423(+)
24                            414-421(+)
25                            679-686(+)
26                            712-724(+)
```

Triplets:

```
7-14-16     23-30(-)          76-88(+)          137-144(-)
7-15-16     23-30(-)          77-89(+)          137-144(-)
7-16-16     23-30(-)          80-87(+)          137-144(-)
8-14-16     39-46(-)          76-88(+)          137-144(-)
8-15-16     39-46(-)          77-89(+)          137-144(-)
8-16-16     39-46(-)          80-87(+)          137-144(-)
************************************************************
```

```
       + ELF-1               - STAT1           - Pax-4
============================================================
1      -1467--1456(+)        -1462--1455(-)    -1469--1458(-)
```

2	-1020--1009(+)	-979--972(-)	-1399--1389(-)
3	34-45(+)	-851--844(-)	-1207--1178(-)
4	492-503(+)	-842--835(-)	-1171--1160(-)
5		-804--797(-)	-1077--1057(-)
6		-679--672(-)	-1003--983(-)
7		-616--609(-)	-971--951(-)
8		-546--539(-)	-922--912(-)
9		-409--402(-)	-691--662(-)
10		-312--305(-)	-690--661(-)
11		-285--278(-)	-684--655(-)
12		-278--271(-)	-674--645(-)
13		-254--247(-)	-654--644(-)
14		-20--13(-)	-616--606(-)
15		-2-5(-)	-603--593(-)
16		27-34(-)	-520--509(-)
17		39-46(-)	-349--320(-)
18		74-81(-)	-299--288(-)
19		84-91(-)	-237--226(-)
20		137-144(-)	-133--122(-)
21		418-425(-)	-116--96(-)
22		471-478(-)	197-207(-)
23		497-504(-)	250-279(-)
24		516-523(-)	521-532(-)
25		687-694(-)	524-535(-)
26		742-749(-)	528-539(-)
27		867-874(-)	558-578(-)
28			584-613(-)
29			646-675(-)
30			718-729(-)
31			842-853(-)

Triplets:

2-2-7	-1020--1009(+)	-979--972(-)	-971--951(-)
2-2-8	-1020--1009(+)	-979--972(-)	-922--912(-)
4-24-25	492-503(+)	516-523(-)	524-535(-)
4-24-26	492-503(+)	516-523(-)	528-539(-)
4-24-27	492-503(+)	516-523(-)	558-578(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1421--1408(+)	-287--280(+)	-1462--1455(-)
2	-841--828(+)	416-423(+)	-979--972(-)
3	-840--827(+)		-851--844(-)
4	-839--826(+)		-842--835(-)
5	-838--825(+)		-804--797(-)
6	-778--765(+)		-679--672(-)
7	-777--764(+)		-616--609(-)
8	-604--591(+)		-546--539(-)
9	-410--397(+)		-409--402(-)
10	-409--396(+)		-312--305(-)
11	-406--393(+)		-285--278(-)
12	-405--392(+)		-278--271(-)
13	-397--384(+)		-254--247(-)
14	-388--375(+)		-20--13(-)
15	-318--305(+)		-2-5(-)
16	-315--302(+)		27-34(-)

17	-309--296(+)	39-46(-)
18	-299--286(+)	74-81(-)
19	-292--279(+)	84-91(-)
20	-291--278(+)	137-144(-)
21	-290--277(+)	418-425(-)
22	-289--276(+)	471-478(-)
23	-283--270(+)	497-504(-)
24	-278--265(+)	516-523(-)
25	-275--262(+)	687-694(-)
26	-184--171(+)	742-749(-)
27	-183--170(+)	867-874(-)
28	-180--167(+)	
29	-179--166(+)	
30	-170--157(+)	
31	-166--153(+)	
32	-151--138(+)	
33	-146--133(+)	
34	-145--132(+)	
35	-140--127(+)	
36	-139--126(+)	
37	-138--125(+)	
38	-137--124(+)	
39	-107--94(+)	
40	-4-9(+)	
41	12-25(+)	
42	16-29(+)	
43	39-52(+)	
44	62-75(+)	
45	63-76(+)	
46	68-81(+)	
47	69-82(+)	
48	70-83(+)	
49	71-84(+)	
50	72-85(+)	
51	73-86(+)	
52	126-139(+)	
53	224-237(+)	
54	348-361(+)	
55	357-370(+)	
56	374-387(+)	
57	375-388(+)	
58	376-389(+)	
59	377-390(+)	
60	399-412(+)	
61	407-420(+)	
62	501-514(+)	
63	510-523(+)	
64	511-524(+)	
65	706-719(+)	
66	707-720(+)	
67	708-721(+)	

Triplets:

15-1-12	-318--305(+)	-287--280(+)	-278--271(-)
15-1-13	-318--305(+)	-287--280(+)	-254--247(-)
16-1-12	-315--302(+)	-287--280(+)	-278--271(-)
16-1-13	-315--302(+)	-287--280(+)	-254--247(-)

```
17-1-12     -309--296(+)      -287--280(+)      -278--271(-)
17-1-13     -309--296(+)      -287--280(+)      -254--247(-)
55-2-22     357-370(+)        416-423(+)        471-478(-)
56-2-22     374-387(+)        416-423(+)        471-478(-)
57-2-22     375-388(+)        416-423(+)        471-478(-)
58-2-22     376-389(+)        416-423(+)        471-478(-)
59-2-22     377-390(+)        416-423(+)        471-478(-)
60-2-22     399-412(+)        416-423(+)        471-478(-)
************************************************************
```

Human GRIA2

```
           - STAT6           + MZF1            - STAT3
============================================================
1      -1469--1462(-)        -1473--1466(+)    -1469--1462(-)
2      -1158--1151(-)        -1008--996(+)     -1368--1361(-)
3      -1000--993(-)         -1004--997(+)     -1315--1308(-)
4      -973--966(-)          -318--311(+)      -1227--1220(-)
5      -851--844(-)          -146--139(+)      -1204--1197(-)
6      182-189(-)            -144--132(+)      -1158--1151(-)
7      271-278(-)            3-15(+)           -1000--993(-)
8      288-295(-)            24-36(+)          -973--966(-)
9      297-304(-)            138-145(+)        -851--844(-)
10     303-310(-)            289-301(+)        -749--742(-)
11     601-608(-)            309-321(+)        -603--596(-)
12                           388-400(+)        -498--491(-)
13                           392-399(+)        182-189(-)
14                           982-994(+)        288-295(-)
15                                             344-351(-)

Triplets:
7-10-15     271-278(-)        289-301(+)        344-351(-)
7-11-15     271-278(-)        309-321(+)        344-351(-)
8-11-15     288-295(-)        309-321(+)        344-351(-)
9-11-15     297-304(-)        309-321(+)        344-351(-)
************************************************************

           + ELF-1           - STAT1           - Pax-4
============================================================
1      -1474--1463(+)        -1469--1462(-)    -1485--1475(-)
2      283-294(+)            -1368--1361(-)    -1297--1287(-)
3      292-303(+)            -1315--1308(-)    -1287--1277(-)
4      298-309(+)            -1227--1220(-)    -1254--1243(-)
5      596-607(+)            -1193--1186(-)    -1227--1217(-)
6                            -1158--1151(-)    -1075--1046(-)
7                            -1026--1019(-)    -1058--1047(-)
8                            -1000--993(-)     -1054--1025(-)
9                            -973--966(-)      -1035--1025(-)
10                           -851--844(-)      -927--917(-)
11                           -810--803(-)      -773--753(-)
12                           -799--792(-)      -718--708(-)
13                           -749--742(-)      -637--617(-)
14                           -603--596(-)      -611--601(-)
15                           -498--491(-)      -603--593(-)
```

16		-290--283(-)	-552--541(-)
17		101-108(-)	-484--474(-)
18		182-189(-)	-429--419(-)
19		271-278(-)	-273--262(-)
20		288-295(-)	-173--162(-)
21		297-304(-)	-163--152(-)
22		303-310(-)	-22--2(-)
23		344-351(-)	29-40(-)
24		601-608(-)	55-66(-)
25		630-637(-)	231-241(-)
26		741-748(-)	259-269(-)
27			315-344(-)
28			462-491(-)
29			465-494(-)
30			471-481(-)
31			482-492(-)
32			487-516(-)
33			503-513(-)
34			507-517(-)
35			657-667(-)
36			666-676(-)
37			860-880(-)
38			869-898(-)
39			957-986(-)

Triplets:

2-21-27	283-294(+)	297-304(-)	315-344(-)
2-22-27	283-294(+)	303-310(-)	315-344(-)
5-25-35	596-607(+)	630-637(-)	657-667(-)
5-25-36	596-607(+)	630-637(-)	666-676(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1406--1393(+)	269-276(+)	-1469--1462(-)
2	-1012--999(+)		-1368--1361(-)
3	-1011--998(+)		-1315--1308(-)
4	-870--857(+)		-1227--1220(-)
5	-830--817(+)		-1193--1186(-)
6	-741--728(+)		-1158--1151(-)
7	-545--532(+)		-1026--1019(-)
8	-468--455(+)		-1000--993(-)
9	-148--135(+)		-973--966(-)
10	-147--134(+)		-851--844(-)
11	-3-10(+)		-810--803(-)
12	-2-11(+)		-799--792(-)
13	18-31(+)		-749--742(-)
14	19-32(+)		-603--596(-)
15	45-58(+)		-498--491(-)
16	136-149(+)		-290--283(-)
17	137-150(+)		101-108(-)
18	178-191(+)		182-189(-)
19	179-192(+)		271-278(-)
20	244-257(+)		288-295(-)
21	277-290(+)		297-304(-)
22	283-296(+)		303-310(-)
23	286-299(+)		344-351(-)

24	292-305(+)		601-608(-)
25	303-316(+)		630-637(-)
26	304-317(+)		741-748(-)
27	305-318(+)		
28	590-603(+)		
29	625-638(+)		
30	735-748(+)		

Triplets:

20-1-20	244-257(+)	269-276(+)	288-295(-)
20-1-21	244-257(+)	269-276(+)	297-304(-)
20-1-22	244-257(+)	269-276(+)	303-310(-)

****

Human GRIA3

	- STAT6	+ MZF1	- STAT3
1	-1435--1428(-)	-1421--1414(+)	-1195--1188(-)
2	-1195--1188(-)	-1328--1316(+)	-736--729(-)
3	-80--73(-)	-1202--1190(+)	-417--410(-)
4	-22--15(-)	-1199--1192(+)	-80--73(-)
5	44-51(-)	-866--859(+)	-22--15(-)
6	100-107(-)	-864--852(+)	44-51(-)
7	709-716(-)	-839--832(+)	100-107(-)
8		-743--731(+)	440-447(-)
9		-88--76(+)	500-507(-)
10		-87--75(+)	709-716(-)
11		-84--77(+)	
12		-77--65(+)	
13		-49--37(+)	
14		23-35(+)	
15		123-130(+)	
16		125-137(+)	
17		493-505(+)	
18		496-503(+)	
19		548-560(+)	
20		816-823(+)	
21		874-886(+)	

Triplets:

3-13-5	-80--73(-)	-49--37(+)	-22--15(-)
4-14-6	-22--15(-)	23-35(+)	44-51(-)

****

	+ ELF-1	- STAT1	- Pax-4
1	-27--16(+)	-1441--1434(-)	-1423--1412(-)
2	39-50(+)	-1435--1428(-)	-1367--1338(-)
3	704-715(+)	-1373--1366(-)	-1179--1169(-)
4		-1328--1321(-)	-1015--986(-)
5		-1253--1246(-)	-1009--980(-)
6		-1183--1176(-)	-988--959(-)
7		-736--729(-)	-934--914(-)

8		-595--588(-)	-829--818(-)
9		-417--410(-)	-616--605(-)
10		-159--152(-)	-574--563(-)
11		-80--73(-)	-397--368(-)
12		-27--20(-)	-318--289(-)
13		-22--15(-)	-314--285(-)
14		95-102(-)	-298--288(-)
15		100-107(-)	-249--220(-)
16		242-249(-)	-240--211(-)
17		440-447(-)	-105--94(-)
18		500-507(-)	-88--77(-)
19		709-716(-)	-70--59(-)
20			-68--57(-)
21			86-97(-)
22			91-102(-)
23			130-159(-)
24			132-161(-)
25			132-143(-)
26			134-163(-)
27			147-157(-)
28			148-177(-)
29			167-177(-)
30			175-195(-)
31			413-424(-)
32			518-529(-)
33			565-594(-)
34			574-585(-)
35			644-655(-)
36			722-733(-)
37			770-781(-)
38			822-833(-)
39			872-883(-)
40			894-905(-)
41			942-952(-)

Triplets:

2-14-23	39-50(+)	95-102(-)	130-159(-)
2-14-24	39-50(+)	95-102(-)	132-161(-)
2-14-25	39-50(+)	95-102(-)	132-143(-)
2-14-26	39-50(+)	95-102(-)	134-163(-)
2-14-27	39-50(+)	95-102(-)	147-157(-)
2-14-28	39-50(+)	95-102(-)	148-177(-)
2-15-23	39-50(+)	100-107(-)	130-159(-)
2-15-24	39-50(+)	100-107(-)	132-161(-)
2-15-25	39-50(+)	100-107(-)	132-143(-)
2-15-26	39-50(+)	100-107(-)	134-163(-)
2-15-27	39-50(+)	100-107(-)	147-157(-)
2-15-28	39-50(+)	100-107(-)	148-177(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1457--1444(+)	-1197--1190(+)	-1441--1434(-)
2	-1446--1433(+)	-82--75(+)	-1435--1428(-)
3	-1362--1349(+)	42-49(+)	-1373--1366(-)
4	-1333--1320(+)		-1328--1321(-)
5	-1332--1319(+)		-1253--1246(-)

6	-1327--1314(+)	-1183--1176(-)
7	-1326--1313(+)	-736--729(-)
8	-1206--1193(+)	-595--588(-)
9	-870--857(+)	-417--410(-)
10	-869--856(+)	-159--152(-)
11	-834--821(+)	-80--73(-)
12	-791--778(+)	-27--20(-)
13	-652--639(+)	-22--15(-)
14	-622--609(+)	95-102(-)
15	-339--326(+)	100-107(-)
16	-338--325(+)	242-249(-)
17	-337--324(+)	440-447(-)
18	-98--85(+)	500-507(-)
19	-94--81(+)	709-716(-)
20	-93--80(+)	
21	-92--79(+)	
22	-87--74(+)	
23	-86--73(+)	
24	-83--70(+)	
25	-82--69(+)	
26	-81--68(+)	
27	-80--67(+)	
28	-54--41(+)	
29	-45--32(+)	
30	-44--31(+)	
31	-33--20(+)	
32	17-30(+)	
33	18-31(+)	
34	33-46(+)	
35	42-55(+)	
36	45-58(+)	
37	48-61(+)	
38	66-79(+)	
39	69-82(+)	
40	89-102(+)	
41	114-127(+)	
42	115-128(+)	
43	116-129(+)	
44	119-132(+)	
45	120-133(+)	
46	121-134(+)	
47	122-135(+)	
48	238-251(+)	
49	239-252(+)	
50	240-253(+)	
51	443-456(+)	
52	482-495(+)	
53	542-555(+)	
54	543-556(+)	
55	544-557(+)	
56	620-633(+)	
57	632-645(+)	
58	882-895(+)	
59	883-896(+)	
60	884-897(+)	

Triplets:

```
18-2-12     -98--85(+)        -82--75(+)        -27--20(-)
32-3-14     17-30(+)          42-49(+)          95-102(-)
32-3-15     17-30(+)          42-49(+)          100-107(-)
33-3-14     18-31(+)          42-49(+)          95-102(-)
33-3-15     18-31(+)          42-49(+)          100-107(-)
************************************************************
```

Human GRIA4

```
           - STAT6           + MZF1            - STAT3
============================================================
1      -1418--1411(-)        97-104(+)         -1286--1279(-)
2      -1286--1279(-)        205-217(+)        -1231--1224(-)
3      -1231--1224(-)        288-300(+)        -916--909(-)
4      -884--877(-)          384-396(+)        -884--877(-)
5      -863--856(-)          415-427(+)        -863--856(-)
6      -473--466(-)          418-425(+)        -459--452(-)
7      20-27(-)              466-473(+)        20-27(-)
8      196-203(-)            541-553(+)        258-265(-)
9      258-265(-)            901-913(+)        295-302(-)
10     295-302(-)            962-969(+)        690-697(-)
11     690-697(-)                              864-871(-)
12     864-871(-)
```

Triplets:

```
8-2-8 196-203(-)        205-217(+)        258-265(-)
************************************************************
```



```
           + ELF-1           - STAT1           - Pax-4
============================================================
1      -1423--1412(+)        -1475--1468(-)    -1467--1438(-)
2      -1236--1225(+)        -1433--1426(-)    -1463--1453(-)
3      -889--878(+)          -1418--1411(-)    -1342--1313(-)
4      253-264(+)            -1298--1291(-)    -1294--1265(-)
5      859-870(+)            -1286--1279(-)    -1260--1250(-)
6                            -1231--1224(-)    -1165--1155(-)
7                            -1225--1218(-)    -1124--1114(-)
8                            -916--909(-)      -942--932(-)
9                            -884--877(-)      -923--894(-)
10                           -863--856(-)      -815--804(-)
11                           -856--849(-)      -718--708(-)
12                           -783--776(-)      -646--635(-)
13                           -591--584(-)      -454--443(-)
14                           -473--466(-)      -432--422(-)
15                           -459--452(-)      -425--415(-)
16                           -442--435(-)      -412--392(-)
17                           -53--46(-)        -356--327(-)
18                           20-27(-)          -338--328(-)
19                           196-203(-)        -198--188(-)
20                           261-268(-)        -58--47(-)
21                           295-302(-)        -52--42(-)
22                           537-544(-)        220-231(-)
23                           658-665(-)        282-293(-)
24                           690-697(-)        291-311(-)
```

25		750-757(-)	347-358(-)
26		864-871(-)	362-373(-)
27		877-884(-)	379-390(-)
28		908-915(-)	500-511(-)
29			777-806(-)
30			781-810(-)
31			818-828(-)

Triplets:

3-10-10	-889--878(+)	-863--856(-)	-815--804(-)
3-11-10	-889--878(+)	-856--849(-)	-815--804(-)
4-21-25	253-264(+)	295-302(-)	347-358(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1429--1416(+)	-865--858(+)	-1475--1468(-)
2	-874--861(+)	256-263(+)	-1433--1426(-)
3	-860--847(+)		-1418--1411(-)
4	-789--776(+)		-1298--1291(-)
5	-6-7(+)		-1286--1279(-)
6	-5-8(+)		-1231--1224(-)
7	0-13(+)		-1225--1218(-)
8	9-22(+)		-916--909(-)
9	17-30(+)		-884--877(-)
10	30-43(+)		-863--856(-)
11	31-44(+)		-856--849(-)
12	34-47(+)		-783--776(-)
13	75-88(+)		-591--584(-)
14	199-212(+)		-473--466(-)
15	200-213(+)		-459--452(-)
16	201-214(+)		-442--435(-)
17	258-271(+)		-53--46(-)
18	261-274(+)		20-27(-)
19	262-275(+)		196-203(-)
20	284-297(+)		261-268(-)
21	411-424(+)		295-302(-)
22	423-436(+)		537-544(-)
23	535-548(+)		658-665(-)
24	536-549(+)		690-697(-)
25	537-550(+)		750-757(-)
26	911-924(+)		864-871(-)
27			877-884(-)
28			908-915(-)

Triplets:

14-2-21	199-212(+)	256-263(+)	295-302(-)
15-2-21	200-213(+)	256-263(+)	295-302(-)
16-2-21	201-214(+)	256-263(+)	295-302(-)

****

Murine GRIA1

	- STAT6	+ MZF1	- STAT3

1	-1219--1212(-)	-1463--1451(+)	-1228--1221(-)
2	-1215--1208(-)	-1262--1250(+)	-1223--1216(-)
3	-1211--1204(-)	-1257--1250(+)	-1219--1212(-)
4	-1207--1200(-)	-1107--1100(+)	-1215--1208(-)
5	-1203--1196(-)	-1105--1093(+)	-1211--1204(-)
6	-1199--1192(-)	-1100--1093(+)	-1207--1200(-)
7	-1195--1188(-)	-1098--1086(+)	-1203--1196(-)
8	-1191--1184(-)	-959--952(+)	-1199--1192(-)
9	-1183--1176(-)	-947--935(+)	-1195--1188(-)
10	-1166--1159(-)	-652--640(+)	-1191--1184(-)
11	-1089--1082(-)	-649--642(+)	-1166--1159(-)
12	-1021--1014(-)	-187--175(+)	-1089--1082(-)
13	-1017--1010(-)	-183--176(+)	-1021--1014(-)
14	-1013--1006(-)	-68--56(+)	-1017--1010(-)
15	-1001--994(-)	152-164(+)	-945--938(-)
16	-945--938(-)	156-163(+)	-844--837(-)
17	-626--619(-)	616-628(+)	-761--754(-)
18	-443--436(-)	908-920(+)	-626--619(-)
19	-380--373(-)	912-919(+)	-557--550(-)
20	-134--127(-)		-443--436(-)
21	-10--3(-)		-380--373(-)
22	67-74(-)		-134--127(-)
23	622-629(-)		-34--27(-)
24			-10--3(-)
25			67-74(-)
26			622-629(-)
27			916-923(-)

Triplets:

13-8-15	-1017--1010(-)	-959--952(+)	-945--938(-)
14-8-15	-1013--1006(-)	-959--952(+)	-945--938(-)
15-8-15	-1001--994(-)	-959--952(+)	-945--938(-)

	+ ELF-1	- STAT1	- Pax-4
1	-1188--1177(+)	-1488--1481(-)	-1393--1373(-)
2	-1026--1015(+)	-1390--1383(-)	-931--902(-)
3	-849--838(+)	-1313--1306(-)	-511--501(-)
4	-631--620(+)	-1295--1288(-)	-451--441(-)
5	-385--374(+)	-1228--1221(-)	-314--285(-)
6	-139--128(+)	-1191--1184(-)	-309--280(-)
7		-1183--1176(-)	-306--277(-)
8		-1176--1169(-)	-304--275(-)
9		-1166--1159(-)	-299--270(-)
10		-1162--1155(-)	-297--268(-)
11		-1158--1151(-)	-296--267(-)
12		-1154--1147(-)	-295--266(-)
13		-1150--1143(-)	-294--265(-)
14		-1146--1139(-)	-130--101(-)
15		-1142--1135(-)	-50--39(-)
16		-1138--1131(-)	-14--3(-)
17		-1134--1127(-)	159-169(-)
18		-1130--1123(-)	397-407(-)
19		-1126--1119(-)	483-512(-)
20		-1122--1115(-)	499-510(-)

21		-1118--1111(-)	501-530(-)
22		-1114--1107(-)	540-551(-)
23		-1110--1103(-)	651-680(-)
24		-1086--1079(-)	746-775(-)
25		-1081--1074(-)	763-773(-)
26		-1077--1070(-)	766-776(-)
27		-1073--1066(-)	781-792(-)
28		-1069--1062(-)	958-969(-)
29		-1065--1058(-)	965-975(-)
30		-1061--1054(-)	
31		-1057--1050(-)	
32		-1053--1046(-)	
33		-1049--1042(-)	
34		-1045--1038(-)	
35		-1041--1034(-)	
36		-1037--1030(-)	
37		-1033--1026(-)	
38		-1029--1022(-)	
39		-1013--1006(-)	
40		-1009--1002(-)	
41		-1001--994(-)	
42		-997--990(-)	
43		-993--986(-)	
44		-989--982(-)	
45		-985--978(-)	
46		-981--974(-)	
47		-977--970(-)	
48		-973--966(-)	
49		-969--962(-)	
50		-945--938(-)	
51		-854--847(-)	
52		-844--837(-)	
53		-634--627(-)	
54		-626--619(-)	
55		-557--550(-)	
56		-456--449(-)	
57		-452--445(-)	
58		-443--436(-)	
59		-339--332(-)	
60		-335--328(-)	
61		-316--309(-)	
62		-310--303(-)	
63		-145--138(-)	
64		-134--127(-)	
65		-10--3(-)	
66		67-74(-)	
67		415-422(-)	
68		595-602(-)	
69		622-629(-)	
70		697-704(-)	
71		774-781(-)	
72		916-923(-)	
73		982-989(-)	

Triplets:

2-44-2	-1026--1015(+)	-989--982(-)	-931--902(-)
2-45-2	-1026--1015(+)	-985--978(-)	-931--902(-)

2-46-2	-1026--1015(+)	-981--974(-)	-931--902(-)
2-47-2	-1026--1015(+)	-977--970(-)	-931--902(-)
2-48-2	-1026--1015(+)	-973--966(-)	-931--902(-)
2-49-2	-1026--1015(+)	-969--962(-)	-931--902(-)
5-59-5	-385--374(+)	-339--332(-)	-314--285(-)
5-59-6	-385--374(+)	-339--332(-)	-309--280(-)
5-59-7	-385--374(+)	-339--332(-)	-306--277(-)
5-59-8	-385--374(+)	-339--332(-)	-304--275(-)
5-59-9	-385--374(+)	-339--332(-)	-299--270(-)
5-59-10	-385--374(+)	-339--332(-)	-297--268(-)
5-59-11	-385--374(+)	-339--332(-)	-296--267(-)
5-59-12	-385--374(+)	-339--332(-)	-295--266(-)
5-59-13	-385--374(+)	-339--332(-)	-294--265(-)
5-60-5	-385--374(+)	-335--328(-)	-314--285(-)
5-60-6	-385--374(+)	-335--328(-)	-309--280(-)
5-60-7	-385--374(+)	-335--328(-)	-306--277(-)
5-60-8	-385--374(+)	-335--328(-)	-304--275(-)
5-60-9	-385--374(+)	-335--328(-)	-299--270(-)
5-60-10	-385--374(+)	-335--328(-)	-297--268(-)
5-60-11	-385--374(+)	-335--328(-)	-296--267(-)
5-60-12	-385--374(+)	-335--328(-)	-295--266(-)
5-60-13	-385--374(+)	-335--328(-)	-294--265(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1478--1465(+)	-1091--1084(+)	-1488--1481(-)
2	-1387--1374(+)	-947--940(+)	-1390--1383(-)
3	-1317--1304(+)	-628--621(+)	-1313--1306(-)
4	-1289--1276(+)	-12--5(+)	-1295--1288(-)
5	-1268--1255(+)		-1228--1221(-)
6	-1267--1254(+)		-1191--1184(-)
7	-1266--1253(+)		-1183--1176(-)
8	-1265--1252(+)		-1176--1169(-)
9	-1264--1251(+)		-1166--1159(-)
10	-1259--1246(+)		-1162--1155(-)
11	-1258--1245(+)		-1158--1151(-)
12	-1255--1242(+)		-1154--1147(-)
13	-1254--1241(+)		-1150--1143(-)
14	-1250--1237(+)		-1146--1139(-)
15	-1246--1233(+)		-1142--1135(-)
16	-1234--1221(+)		-1138--1131(-)
17	-1230--1217(+)		-1134--1127(-)
18	-1226--1213(+)		-1130--1123(-)
19	-1222--1209(+)		-1126--1119(-)
20	-1218--1205(+)		-1122--1115(-)
21	-1214--1201(+)		-1118--1111(-)
22	-1210--1197(+)		-1114--1107(-)
23	-1206--1193(+)		-1110--1103(-)
24	-1202--1189(+)		-1086--1079(-)
25	-1197--1184(+)		-1081--1074(-)
26	-1194--1181(+)		-1077--1070(-)
27	-1181--1168(+)		-1073--1066(-)
28	-1178--1165(+)		-1069--1062(-)
29	-1177--1164(+)		-1065--1058(-)
30	-1116--1103(+)		-1061--1054(-)
31	-1115--1102(+)		-1057--1050(-)

32	-1114--1101(+)	-1053--1046(-)
33	-1111--1098(+)	-1049--1042(-)
34	-1110--1097(+)	-1045--1038(-)
35	-1109--1096(+)	-1041--1034(-)
36	-1108--1095(+)	-1037--1030(-)
37	-1107--1094(+)	-1033--1026(-)
38	-1104--1091(+)	-1029--1022(-)
39	-1103--1090(+)	-1013--1006(-)
40	-1100--1087(+)	-1009--1002(-)
41	-1032--1019(+)	-1001--994(-)
42	-1028--1015(+)	-997--990(-)
43	-1024--1011(+)	-993--986(-)
44	-1012--999(+)	-989--982(-)
45	-968--955(+)	-985--978(-)
46	-967--954(+)	-981--974(-)
47	-966--953(+)	-977--970(-)
48	-963--950(+)	-973--966(-)
49	-960--947(+)	-969--962(-)
50	-959--946(+)	-945--938(-)
51	-956--943(+)	-854--847(-)
52	-952--939(+)	-844--837(-)
53	-951--938(+)	-634--627(-)
54	-749--736(+)	-626--619(-)
55	-658--645(+)	-557--550(-)
56	-657--644(+)	-456--449(-)
57	-656--643(+)	-452--445(-)
58	-637--624(+)	-443--436(-)
59	-472--459(+)	-339--332(-)
60	-455--442(+)	-335--328(-)
61	-454--441(+)	-316--309(-)
62	-196--183(+)	-310--303(-)
63	-192--179(+)	-145--138(-)
64	-191--178(+)	-134--127(-)
65	-190--177(+)	-10--3(-)
66	-145--132(+)	67-74(-)
67	-74--61(+)	415-422(-)
68	-60--47(+)	595-602(-)
69	32-45(+)	622-629(-)
70	344-357(+)	697-704(-)
71	611-624(+)	774-781(-)
72	617-630(+)	916-923(-)
73	843-856(+)	982-989(-)
74	905-918(+)	

Triplets:

30-1-25	-1116--1103(+)	-1091--1084(+)	-1081--1074(-)
30-1-26	-1116--1103(+)	-1091--1084(+)	-1077--1070(-)
30-1-27	-1116--1103(+)	-1091--1084(+)	-1073--1066(-)
30-1-28	-1116--1103(+)	-1091--1084(+)	-1069--1062(-)
30-1-29	-1116--1103(+)	-1091--1084(+)	-1065--1058(-)
30-1-30	-1116--1103(+)	-1091--1084(+)	-1061--1054(-)
30-1-31	-1116--1103(+)	-1091--1084(+)	-1057--1050(-)
30-1-32	-1116--1103(+)	-1091--1084(+)	-1053--1046(-)
30-1-33	-1116--1103(+)	-1091--1084(+)	-1049--1042(-)
30-1-34	-1116--1103(+)	-1091--1084(+)	-1045--1038(-)
30-1-35	-1116--1103(+)	-1091--1084(+)	-1041--1034(-)
30-1-36	-1116--1103(+)	-1091--1084(+)	-1037--1030(-)

30-1-37	-1116--1103(+)	-1091--1084(+)	-1033--1026(-)
31-1-25	-1115--1102(+)	-1091--1084(+)	-1081--1074(-)
31-1-26	-1115--1102(+)	-1091--1084(+)	-1077--1070(-)
31-1-27	-1115--1102(+)	-1091--1084(+)	-1073--1066(-)
31-1-28	-1115--1102(+)	-1091--1084(+)	-1069--1062(-)
31-1-29	-1115--1102(+)	-1091--1084(+)	-1065--1058(-)
31-1-30	-1115--1102(+)	-1091--1084(+)	-1061--1054(-)
31-1-31	-1115--1102(+)	-1091--1084(+)	-1057--1050(-)
31-1-32	-1115--1102(+)	-1091--1084(+)	-1053--1046(-)
31-1-33	-1115--1102(+)	-1091--1084(+)	-1049--1042(-)
31-1-34	-1115--1102(+)	-1091--1084(+)	-1045--1038(-)
31-1-35	-1115--1102(+)	-1091--1084(+)	-1041--1034(-)
31-1-36	-1115--1102(+)	-1091--1084(+)	-1037--1030(-)
31-1-37	-1115--1102(+)	-1091--1084(+)	-1033--1026(-)
32-1-25	-1114--1101(+)	-1091--1084(+)	-1081--1074(-)
32-1-26	-1114--1101(+)	-1091--1084(+)	-1077--1070(-)
32-1-27	-1114--1101(+)	-1091--1084(+)	-1073--1066(-)
32-1-28	-1114--1101(+)	-1091--1084(+)	-1069--1062(-)
32-1-29	-1114--1101(+)	-1091--1084(+)	-1065--1058(-)
32-1-30	-1114--1101(+)	-1091--1084(+)	-1061--1054(-)
32-1-31	-1114--1101(+)	-1091--1084(+)	-1057--1050(-)
32-1-32	-1114--1101(+)	-1091--1084(+)	-1053--1046(-)
32-1-33	-1114--1101(+)	-1091--1084(+)	-1049--1042(-)
32-1-34	-1114--1101(+)	-1091--1084(+)	-1045--1038(-)
32-1-35	-1114--1101(+)	-1091--1084(+)	-1041--1034(-)
32-1-36	-1114--1101(+)	-1091--1084(+)	-1037--1030(-)
32-1-37	-1114--1101(+)	-1091--1084(+)	-1033--1026(-)
33-1-25	-1111--1098(+)	-1091--1084(+)	-1081--1074(-)
33-1-26	-1111--1098(+)	-1091--1084(+)	-1077--1070(-)
33-1-27	-1111--1098(+)	-1091--1084(+)	-1073--1066(-)
33-1-28	-1111--1098(+)	-1091--1084(+)	-1069--1062(-)
33-1-29	-1111--1098(+)	-1091--1084(+)	-1065--1058(-)
33-1-30	-1111--1098(+)	-1091--1084(+)	-1061--1054(-)
33-1-31	-1111--1098(+)	-1091--1084(+)	-1057--1050(-)
33-1-32	-1111--1098(+)	-1091--1084(+)	-1053--1046(-)
33-1-33	-1111--1098(+)	-1091--1084(+)	-1049--1042(-)
33-1-34	-1111--1098(+)	-1091--1084(+)	-1045--1038(-)
33-1-35	-1111--1098(+)	-1091--1084(+)	-1041--1034(-)
33-1-36	-1111--1098(+)	-1091--1084(+)	-1037--1030(-)
33-1-37	-1111--1098(+)	-1091--1084(+)	-1033--1026(-)
34-1-25	-1110--1097(+)	-1091--1084(+)	-1081--1074(-)
34-1-26	-1110--1097(+)	-1091--1084(+)	-1077--1070(-)
34-1-27	-1110--1097(+)	-1091--1084(+)	-1073--1066(-)
34-1-28	-1110--1097(+)	-1091--1084(+)	-1069--1062(-)
34-1-29	-1110--1097(+)	-1091--1084(+)	-1065--1058(-)
34-1-30	-1110--1097(+)	-1091--1084(+)	-1061--1054(-)
34-1-31	-1110--1097(+)	-1091--1084(+)	-1057--1050(-)
34-1-32	-1110--1097(+)	-1091--1084(+)	-1053--1046(-)
34-1-33	-1110--1097(+)	-1091--1084(+)	-1049--1042(-)
34-1-34	-1110--1097(+)	-1091--1084(+)	-1045--1038(-)
34-1-35	-1110--1097(+)	-1091--1084(+)	-1041--1034(-)
34-1-36	-1110--1097(+)	-1091--1084(+)	-1037--1030(-)
34-1-37	-1110--1097(+)	-1091--1084(+)	-1033--1026(-)
35-1-25	-1109--1096(+)	-1091--1084(+)	-1081--1074(-)
35-1-26	-1109--1096(+)	-1091--1084(+)	-1077--1070(-)
35-1-27	-1109--1096(+)	-1091--1084(+)	-1073--1066(-)
35-1-28	-1109--1096(+)	-1091--1084(+)	-1069--1062(-)

35-1-29	-1109--1096(+)	-1091--1084(+)	-1065--1058(-)
35-1-30	-1109--1096(+)	-1091--1084(+)	-1061--1054(-)
35-1-31	-1109--1096(+)	-1091--1084(+)	-1057--1050(-)
35-1-32	-1109--1096(+)	-1091--1084(+)	-1053--1046(-)
35-1-33	-1109--1096(+)	-1091--1084(+)	-1049--1042(-)
35-1-34	-1109--1096(+)	-1091--1084(+)	-1045--1038(-)
35-1-35	-1109--1096(+)	-1091--1084(+)	-1041--1034(-)
35-1-36	-1109--1096(+)	-1091--1084(+)	-1037--1030(-)
35-1-37	-1109--1096(+)	-1091--1084(+)	-1033--1026(-)
36-1-25	-1108--1095(+)	-1091--1084(+)	-1081--1074(-)
36-1-26	-1108--1095(+)	-1091--1084(+)	-1077--1070(-)
36-1-27	-1108--1095(+)	-1091--1084(+)	-1073--1066(-)
36-1-28	-1108--1095(+)	-1091--1084(+)	-1069--1062(-)
36-1-29	-1108--1095(+)	-1091--1084(+)	-1065--1058(-)
36-1-30	-1108--1095(+)	-1091--1084(+)	-1061--1054(-)
36-1-31	-1108--1095(+)	-1091--1084(+)	-1057--1050(-)
36-1-32	-1108--1095(+)	-1091--1084(+)	-1053--1046(-)
36-1-33	-1108--1095(+)	-1091--1084(+)	-1049--1042(-)
36-1-34	-1108--1095(+)	-1091--1084(+)	-1045--1038(-)
36-1-35	-1108--1095(+)	-1091--1084(+)	-1041--1034(-)
36-1-36	-1108--1095(+)	-1091--1084(+)	-1037--1030(-)
36-1-37	-1108--1095(+)	-1091--1084(+)	-1033--1026(-)
37-1-25	-1107--1094(+)	-1091--1084(+)	-1081--1074(-)
37-1-26	-1107--1094(+)	-1091--1084(+)	-1077--1070(-)
37-1-27	-1107--1094(+)	-1091--1084(+)	-1073--1066(-)
37-1-28	-1107--1094(+)	-1091--1084(+)	-1069--1062(-)
37-1-29	-1107--1094(+)	-1091--1084(+)	-1065--1058(-)
37-1-30	-1107--1094(+)	-1091--1084(+)	-1061--1054(-)
37-1-31	-1107--1094(+)	-1091--1084(+)	-1057--1050(-)
37-1-32	-1107--1094(+)	-1091--1084(+)	-1053--1046(-)
37-1-33	-1107--1094(+)	-1091--1084(+)	-1049--1042(-)
37-1-34	-1107--1094(+)	-1091--1084(+)	-1045--1038(-)
37-1-35	-1107--1094(+)	-1091--1084(+)	-1041--1034(-)
37-1-36	-1107--1094(+)	-1091--1084(+)	-1037--1030(-)
37-1-37	-1107--1094(+)	-1091--1084(+)	-1033--1026(-)

Murine GRIA2

	- STAT6	+ MZF1	- STAT3
1	-1498--1491(-)	-1493--1486(+)	-1498--1491(-)
2	-1422--1415(-)	-1491--1479(+)	-1422--1415(-)
3	-497--490(-)	-1468--1456(+)	-497--490(-)
4	361-368(-)	-699--692(+)	-335--328(-)
5	378-385(-)	-631--619(+)	-277--270(-)
6	387-394(-)	-53--46(+)	387-394(-)
7	393-400(-)	-51--39(+)	393-400(-)
8	721-728(-)	97-109(+)	432-439(-)
9	801-808(-)	118-130(+)	721-728(-)
10		396-408(+)	801-808(-)
11		399-406(+)	
12		476-488(+)	
13		480-487(+)	
14		714-726(+)	

15 717-724(+)

Triplets:

1-3-2	-1498--1491(-)	-1468--1456(+)	-1422--1415(-)
4-10-8	361-368(-)	396-408(+)	432-439(-)
4-11-8	361-368(-)	399-406(+)	432-439(-)
5-10-8	378-385(-)	396-408(+)	432-439(-)
5-11-8	378-385(-)	399-406(+)	432-439(-)
6-10-8	387-394(-)	396-408(+)	432-439(-)
6-11-8	387-394(-)	399-406(+)	432-439(-)

****

	+ ELF-1	- STAT1	- Pax-4
1	-1427--1416(+)	-1498--1491(-)	-1461--1450(-)
2	-502--491(+)	-1484--1477(-)	-1420--1391(-)
3	373-384(+)	-1422--1415(-)	-1413--1384(-)
4	382-393(+)	-1364--1357(-)	-1369--1340(-)
5		-1143--1136(-)	-1349--1339(-)
6		-863--856(-)	-1339--1329(-)
7		-771--764(-)	-1276--1247(-)
8		-332--325(-)	-1148--1119(-)
9		-277--270(-)	-1087--1077(-)
10		361-368(-)	-952--942(-)
11		378-385(-)	-945--935(-)
12		387-394(-)	-874--864(-)
13		393-400(-)	-871--861(-)
14		432-439(-)	-574--545(-)
15		644-651(-)	-475--446(-)
16		689-696(-)	-473--444(-)
17		721-728(-)	-411--382(-)
18		801-808(-)	-312--302(-)
19			-276--266(-)
20			-70--59(-)
21			-45--34(-)
22			-41--30(-)
23			85-96(-)
24			123-134(-)
25			147-158(-)
26			301-330(-)
27			349-359(-)
28			385-414(-)
29			401-430(-)
30			403-432(-)
31			550-579(-)
32			553-582(-)
33			559-569(-)
34			570-580(-)
35			575-604(-)
36			591-601(-)
37			595-605(-)
38			833-843(-)
39			963-974(-)

Triplets:

3-12-29	373-384(+)	387-394(-)	401-430(-)
3-12-30	373-384(+)	387-394(-)	403-432(-)

3-13-29	373-384(+)	393-400(-)	401-430(-)
3-13-30	373-384(+)	393-400(-)	403-432(-)

	+ GKLF	+ PU.1	- STAT1
1	-1500--1487(+)	359-366(+)	-1498--1491(-)
2	-1497--1484(+)		-1484--1477(-)
3	-1472--1459(+)		-1422--1415(-)
4	-1471--1458(+)		-1364--1357(-)
5	-1468--1455(+)		-1143--1136(-)
6	-1339--1326(+)		-863--856(-)
7	-1050--1037(+)		-771--764(-)
8	-848--835(+)		-332--325(-)
9	-706--693(+)		-277--270(-)
10	-355--342(+)		361-368(-)
11	-354--341(+)		378-385(-)
12	-207--194(+)		387-394(-)
13	-55--42(+)		393-400(-)
14	-51--38(+)		432-439(-)
15	91-104(+)		644-651(-)
16	92-105(+)		689-696(-)
17	113-126(+)		721-728(-)
18	137-150(+)		801-808(-)
19	221-234(+)		
20	226-239(+)		
21	334-347(+)		
22	373-386(+)		
23	376-389(+)		
24	381-394(+)		
25	382-395(+)		
26	390-403(+)		
27	391-404(+)		
28	392-405(+)		
29	709-722(+)		
30	710-723(+)		
31	715-728(+)		
32	820-833(+)		

Triplets:

21-1-11	334-347(+)	359-366(+)	378-385(-)
21-1-12	334-347(+)	359-366(+)	387-394(-)
21-1-13	334-347(+)	359-366(+)	393-400(-)

Murine GRIA3

	- STAT6	+ MZF1	- STAT3
1	-897--890(-)	-1424--1417(+)	-1416--1409(-)
2	-209--202(-)	-1422--1410(+)	-1125--1118(-)
3	-194--187(-)	-1373--1361(+)	-1016--1009(-)
4	-169--162(-)	-1167--1155(+)	-897--890(-)
5	-102--95(-)	-1163--1156(+)	-702--695(-)
6	341-348(-)	-1161--1149(+)	-611--604(-)

7	497-504(-)	-1136--1129(+)	-316--309(-)
8	642-649(-)	-1134--1122(+)	-209--202(-)
9	673-680(-)	-1129--1122(+)	-194--187(-)
10	708-715(-)	-1023--1011(+)	-169--162(-)
11	934-941(-)	-619--607(+)	-102--95(-)
12		-615--608(+)	341-348(-)
13		-320--313(+)	497-504(-)
14		-247--235(+)	614-621(-)
15		-223--211(+)	642-649(-)
16		-218--211(+)	673-680(-)
17		-186--174(+)	708-715(-)
18		-78--66(+)	775-782(-)
19		-75--68(+)	887-894(-)
20		-73--61(+)	934-941(-)
21		160-172(+)	
22		248-255(+)	
23		263-275(+)	
24		268-275(+)	
25		316-328(+)	
26		320-327(+)	
27		388-400(+)	
28		661-673(+)	
29		860-872(+)	
30		863-870(+)	
31		880-892(+)	
32		883-890(+)	
33		928-940(+)	

Triplets:

2-17-10	-209--202(-)	-186--174(+)	-169--162(-)
3-17-10	-194--187(-)	-186--174(+)	-169--162(-)
8-28-17	642-649(-)	661-673(+)	708-715(-)

****

	+ ELF-1	- STAT1	- Pax-4
1	-902--891(+)	-1416--1409(-)	-1461--1450(-)
2	-214--203(+)	-1125--1118(-)	-1455--1444(-)
3	-199--188(+)	-1016--1009(-)	-1450--1440(-)
4	-174--163(+)	-897--890(-)	-1427--1416(-)
5	637-648(+)	-880--873(-)	-1288--1259(-)
6		-611--604(-)	-1118--1107(-)
7		-330--323(-)	-972--952(-)
8		-316--309(-)	-618--589(-)
9		-209--202(-)	-611--582(-)
10		-194--187(-)	-600--571(-)
11		-169--162(-)	-599--570(-)
12		-107--100(-)	-591--581(-)
13		-102--95(-)	-566--537(-)
14		45-52(-)	-562--533(-)
15		187-194(-)	-512--502(-)
16		341-348(-)	-276--265(-)
17		497-504(-)	-259--248(-)
18		608-615(-)	-240--229(-)
19		614-621(-)	-238--227(-)
20		625-632(-)	-116--105(-)
21		673-680(-)	-111--100(-)

22		708-715(-)	-68--39(-)
23		775-782(-)	-66--37(-)
24		891-898(-)	-66--55(-)
25		934-941(-)	-64--35(-)
26			-62--33(-)
27			-51--41(-)
28			-50--21(-)
29			-31--21(-)
30			357-386(-)
31			366-377(-)
32			510-521(-)
33			533-553(-)
34			668-679(-)
35			681-692(-)
36			781-810(-)
37			823-833(-)
38			892-902(-)
39			960-971(-)
40			966-977(-)

Triplets:

2-11-20	-214--203(+)	-169--162(-)	-116--105(-)
2-11-21	-214--203(+)	-169--162(-)	-111--100(-)
3-11-20	-199--188(+)	-169--162(-)	-116--105(-)
3-11-21	-199--188(+)	-169--162(-)	-111--100(-)
5-21-35	637-648(+)	673-680(-)	681-692(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1484--1471(+)	-211--204(+)	-1416--1409(-)
2	-1480--1467(+)	-196--189(+)	-1125--1118(-)
3	-1471--1458(+)	-171--164(+)	-1016--1009(-)
4	-1427--1414(+)	885-892(+)	-897--890(-)
5	-1379--1366(+)		-880--873(-)
6	-1378--1365(+)		-611--604(-)
7	-1377--1364(+)		-330--323(-)
8	-1166--1153(+)		-316--309(-)
9	-1123--1110(+)		-209--202(-)
10	-901--888(+)		-194--187(-)
11	-269--256(+)		-169--162(-)
12	-265--252(+)		-107--100(-)
13	-264--251(+)		-102--95(-)
14	-263--250(+)		45-52(-)
15	-251--238(+)		187-194(-)
16	-228--215(+)		341-348(-)
17	-227--214(+)		497-504(-)
18	-220--207(+)		608-615(-)
19	-212--199(+)		614-621(-)
20	-205--192(+)		625-632(-)
21	-197--184(+)		673-680(-)
22	-192--179(+)		708-715(-)
23	-191--178(+)		775-782(-)
24	-180--167(+)		891-898(-)
25	-154--141(+)		934-941(-)
26	-151--138(+)		
27	-148--135(+)		

28	-145--132(+)
29	-113--100(+)
30	-82--69(+)
31	-79--66(+)
32	-78--65(+)
33	-77--64(+)
34	41-54(+)
35	42-55(+)
36	43-56(+)
37	156-169(+)
38	157-170(+)
39	260-273(+)
40	330-343(+)
41	334-347(+)
42	335-348(+)
43	336-349(+)
44	389-402(+)
45	390-403(+)
46	409-422(+)
47	603-616(+)
48	657-670(+)
49	658-671(+)
50	662-675(+)
51	670-683(+)
52	688-701(+)
53	769-782(+)
54	865-878(+)
55	875-888(+)
56	955-968(+)

Triplets:

11-1-10	-269--256(+)	-211--204(+)	-194--187(-)
11-1-11	-269--256(+)	-211--204(+)	-169--162(-)
12-1-10	-265--252(+)	-211--204(+)	-194--187(-)
12-1-11	-265--252(+)	-211--204(+)	-169--162(-)
13-1-10	-264--251(+)	-211--204(+)	-194--187(-)
13-1-11	-264--251(+)	-211--204(+)	-169--162(-)
14-1-10	-263--250(+)	-211--204(+)	-194--187(-)
14-1-11	-263--250(+)	-211--204(+)	-169--162(-)
15-1-10	-251--238(+)	-211--204(+)	-194--187(-)
15-1-11	-251--238(+)	-211--204(+)	-169--162(-)
15-2-11	-251--238(+)	-196--189(+)	-169--162(-)
16-1-10	-228--215(+)	-211--204(+)	-194--187(-)
16-1-11	-228--215(+)	-211--204(+)	-169--162(-)
16-2-11	-228--215(+)	-196--189(+)	-169--162(-)
17-1-10	-227--214(+)	-211--204(+)	-194--187(-)
17-1-11	-227--214(+)	-211--204(+)	-169--162(-)
17-2-11	-227--214(+)	-196--189(+)	-169--162(-)
18-2-11	-220--207(+)	-196--189(+)	-169--162(-)
19-2-11	-212--199(+)	-196--189(+)	-169--162(-)
54-4-25	865-878(+)	885-892(+)	934-941(-)

****

Murine GRIA4

	- STAT6	+ MZF1	- STAT3
1	-1262--1255(-)	-166--159(+)	-1262--1255(-)
2	-740--733(-)	-151--144(+)	-740--733(-)
3	-312--305(-)	-9-3(+)	-312--305(-)
4	-223--216(-)	45-57(+)	-223--216(-)
5	47-54(-)	49-56(+)	36-43(-)
6	150-157(-)	143-155(+)	47-54(-)
7	458-465(-)	146-153(+)	150-157(-)
8	492-499(-)	165-177(+)	156-163(-)
9	806-813(-)	172-184(+)	458-465(-)
10	981-988(-)	179-191(+)	492-499(-)
11		186-198(+)	806-813(-)
12		473-480(+)	852-859(-)
13		486-498(+)	964-971(-)
14		687-699(+)	981-988(-)
15		748-755(+)	
16		802-809(+)	

Triplets:

7-12-10	458-465(-)	473-480(+)	492-499(-)

****

Query 1: + ELF-1		- STAT1	- Pax-4
1	-745--734(+)	-1327--1320(-)	-1421--1411(-)
2	42-53(+)	-1262--1255(-)	-1406--1396(-)
3	976-987(+)	-1102--1095(-)	-1299--1270(-)
4		-1016--1009(-)	-1283--1273(-)
5		-740--733(-)	-1163--1134(-)
6		-709--702(-)	-1145--1116(-)
7		-338--331(-)	-1142--1113(-)
8		-312--305(-)	-955--945(-)
9		-223--216(-)	-722--712(-)
10		-26--19(-)	-707--678(-)
11		47-54(-)	-678--668(-)
12		150-157(-)	-658--648(-)
13		156-163(-)	-651--641(-)
14		426-433(-)	-478--468(-)
15		458-465(-)	-460--440(-)
16		492-499(-)	-388--368(-)
17		638-645(-)	-343--332(-)
18		662-669(-)	-337--327(-)
19		806-813(-)	79-99(-)
20		852-859(-)	103-123(-)
21		987-994(-)	132-143(-)
22			185-196(-)
23			198-209(-)
24			288-299(-)
25			305-316(-)
26			640-650(-)
27			713-742(-)
28			834-844(-)
29			894-905(-)

Triplets:

1-6-11	-745--734(+)	-709--702(-)	-678--668(-)
1-6-12	-745--734(+)	-709--702(-)	-658--648(-)
1-6-13	-745--734(+)	-709--702(-)	-651--641(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-751--738(+)	-1264--1257(+)	-1327--1320(-)
2	-229--216(+)	-225--218(+)	-1262--1255(-)
3	-228--215(+)	34-41(+)	-1102--1095(-)
4	-223--210(+)	45-52(+)	-1016--1009(-)
5	-155--142(+)	804-811(+)	-740--733(-)
6	-58--45(+)	979-986(+)	-709--702(-)
7	-26--13(+)		-338--331(-)
8	40-53(+)		-312--305(-)
9	41-54(+)		-223--216(-)
10	42-55(+)		-26--19(-)
11	137-150(+)		47-54(-)
12	144-157(+)		150-157(-)
13	145-158(+)		156-163(-)
14	150-163(+)		426-433(-)
15	151-164(+)		458-465(-)
16	154-167(+)		492-499(-)
17	159-172(+)		638-645(-)
18	166-179(+)		662-669(-)
19	167-180(+)		806-813(-)
20	168-181(+)		852-859(-)
21	173-186(+)		987-994(-)
22	174-187(+)		
23	175-188(+)		
24	181-194(+)		
25	182-195(+)		
26	310-323(+)		
27	311-324(+)		
28	320-333(+)		
29	321-334(+)		
30	660-673(+)		
31	681-694(+)		
32	697-710(+)		
33	793-806(+)		
34	794-807(+)		
35	795-808(+)		
36	801-814(+)		
37	970-983(+)		
38	974-987(+)		

Triplets:

7-3-11	-26--13(+)	34-41(+)	47-54(-)

****

Rat GRIA1

	- STAT6	+ MZF1	- STAT3
1	-1032--1025(-)	-727--715(+)	-1430--1423(-)
2	-325--318(-)	-721--709(+)	-1421--1414(-)
3	-280--273(-)	-720--708(+)	-1032--1025(-)
4	-224--217(-)	-716--709(+)	-536--529(-)
5	152-159(-)	-421--414(+)	-523--516(-)
6	160-167(-)	-356--344(+)	-499--492(-)
7	164-171(-)	-352--345(+)	-325--318(-)
8	189-196(-)	-350--338(+)	-280--273(-)
9	237-244(-)	-319--312(+)	-224--217(-)
10	526-533(-)	-312--300(+)	-166--159(-)
11	549-556(-)	-308--301(+)	148-155(-)
12	593-600(-)	-284--277(+)	152-159(-)
13	702-709(-)	-232--220(+)	160-167(-)
14	890-897(-)	-228--221(+)	180-187(-)
15	909-916(-)	-199--187(+)	189-196(-)
16		-183--176(+)	237-244(-)
17		-109--97(+)	289-296(-)
18		-34--22(+)	526-533(-)
19		21-33(+)	549-556(-)
20		25-32(+)	593-600(-)
21		226-238(+)	702-709(-)
22		229-236(+)	890-897(-)
23		381-393(+)	909-916(-)
24		393-400(+)	
25		406-413(+)	
26		455-467(+)	
27		458-465(+)	
28		578-590(+)	
29		586-598(+)	
30		589-596(+)	
31		696-708(+)	
32		853-865(+)	
33		857-864(+)	
34		881-888(+)	
35		883-895(+)	
36		886-893(+)	

Triplets:

2-10-8	-325--318(-)	-312--300(+)	-280--273(-)
2-11-8	-325--318(-)	-308--301(+)	-280--273(-)
4-15-10	-224--217(-)	-199--187(+)	-166--159(-)
4-16-10	-224--217(-)	-183--176(+)	-166--159(-)
8-21-17	189-196(-)	226-238(+)	289-296(-)
8-22-16	189-196(-)	229-236(+)	237-244(-)
10-28-20	526-533(-)	578-590(+)	593-600(-)
11-28-20	549-556(-)	578-590(+)	593-600(-)

****

	+ ELF-1	- STAT1	- Pax-4
1	-1037--1026(+)	-1471--1464(-)	-1494--1484(-)
2	-330--319(+)	-1421--1414(-)	-1287--1267(-)
3	-171--160(+)	-1032--1025(-)	-1094--1083(-)
4	184-195(+)	-863--856(-)	-594--565(-)
5	521-532(+)	-733--726(-)	-574--564(-)

6		-536--529(-)	-536--526(-)
7		-499--492(-)	-523--513(-)
8		-464--457(-)	-503--492(-)
9		-280--273(-)	-322--311(-)
10		-224--217(-)	-213--202(-)
11		-166--159(-)	-193--182(-)
12		-41--34(-)	-29--18(-)
13		137-144(-)	40-60(-)
14		152-159(-)	290-319(-)
15		164-171(-)	344-355(-)
16		180-187(-)	349-359(-)
17		192-199(-)	410-439(-)
18		237-244(-)	417-446(-)
19		289-296(-)	646-666(-)
20		526-533(-)	653-673(-)
21		549-556(-)	714-734(-)
22		593-600(-)	801-812(-)
23		673-680(-)	914-925(-)
24		702-709(-)	948-977(-)
25		890-897(-)	953-982(-)
26			954-983(-)
27			961-990(-)
28			964-993(-)
29			966-995(-)
30			969-998(-)

Triplets:

4-18-14	184-195(+)	237-244(-)	290-319(-)

****

	+ GKLF	+ PU.1	- STAT1
1	-1377--1364(+)	-501--494(+)	-1471--1464(-)
2	-1376--1363(+)	-327--320(+)	-1421--1414(-)
3	-737--724(+)	158-165(+)	-1032--1025(-)
4	-736--723(+)	524-531(+)	-863--856(-)
5	-733--720(+)	591-598(+)	-733--726(-)
6	-732--719(+)		-536--529(-)
7	-731--718(+)		-499--492(-)
8	-730--717(+)		-464--457(-)
9	-729--716(+)		-280--273(-)
10	-728--715(+)		-224--217(-)
11	-727--714(+)		-166--159(-)
12	-726--713(+)		-41--34(-)
13	-725--712(+)		137-144(-)
14	-724--711(+)		152-159(-)
15	-723--710(+)		164-171(-)
16	-662--649(+)		180-187(-)
17	-634--621(+)		192-199(-)
18	-524--511(+)		237-244(-)
19	-355--342(+)		289-296(-)
20	-354--341(+)		526-533(-)
21	-349--336(+)		549-556(-)
22	-336--323(+)		593-600(-)
23	-332--319(+)		673-680(-)
24	-328--315(+)		702-709(-)
25	-327--314(+)		890-897(-)

26 -326--313(+)
27 -223--210(+)
28 -218--205(+)
29 -209--196(+)
30 -205--192(+)
31 -204--191(+)
32 -203--190(+)
33 -114--101(+)
34 -44--31(+)
35 -43--30(+)
36 -39--26(+)
37 -30--17(+)
38 -23--10(+)
39 -4-9(+)
40 -1-12(+)
41 9-22(+)
42 10-23(+)
43 15-28(+)
44 16-29(+)
45 17-30(+)
46 18-31(+)
47 49-62(+)
48 145-158(+)
49 146-159(+)
50 149-162(+)
51 153-166(+)
52 169-182(+)
53 192-205(+)
54 217-230(+)
55 220-233(+)
56 221-234(+)
57 222-235(+)
58 225-238(+)
59 226-239(+)
60 278-291(+)
61 376-389(+)
62 377-390(+)
63 390-403(+)
64 512-525(+)
65 525-538(+)
66 544-557(+)
67 545-558(+)
68 546-559(+)
69 573-586(+)
70 582-595(+)
71 795-808(+)
72 879-892(+)

Triplets:

18-1-8	-524--511(+)	-501--494(+)	-464--457(-)
19-2-9	-355--342(+)	-327--320(+)	-280--273(-)
20-2-9	-354--341(+)	-327--320(+)	-280--273(-)
21-2-9	-349--336(+)	-327--320(+)	-280--273(-)

****

LITERATURE CITED

Advani RJ, Bae HR, Bock JB, Chao DS, Doung YC, Prekeris R, Yoo JS, Scheller RH (1998) Seven novel mammalian SNARE proteins localize to distinct membrane compartments. *J Biol Chem.* **273**(17): 10317-10324.

Aittomaki S, Yang J, Scott EW, Simon MC, Silvennoinen O (2004) Molecular basis of Stat1 and PU.1 cooperation in cytokine-induced Fcgamma receptor I promoter activation. *Int Immunol.* **16**(2): 265-274.

Albright SR, Tjian R (2000) TAFs revisited: more data reveal new twists and confirm old ideas. *Gene.* **242**(1-2):1-13.

Alfarano C, Andrade CE, Anthony K, Bahroos N, Bajec M, Bantoft K, Betel D, Bobechko B, Boutilier K, Burgess E, Buzadzija K, Cavero R, D'Abreo C, Donaldson I, Dorairajoo D, Dumontier MJ, Dumontier MR, Earles V, Farrall R, Feldman H, Garderman E, Gong Y, Gonzaga R, Grytsan V, Gryz E, Gu V, Haldorsen E, Halupa A, Haw R, Hrvojic A, Hurrell L, Isserlin R, Jack F, Juma F, Khan A, Kon T, Konopinsky S, Le V, Lee E, Ling S, Magidin M, Moniakis J, Montojo J, Moore S, Muskat B, Ng I, Paraiso JP, Parker B, Pintilie G, Pirone R, Salama JJ, Sgro S, Shan T, Shu Y, Siew J, Skinner D, Snyder K, Stasiuk R, Strumpf D, Tuekam B, Tao S, Wang Z, White M, Willis R, Wolting C, Wong S, Wrong A, Xin C, Yao R, Yates B, Zhang S, Zheng K, Pawson T, Ouellette BF, Hogue CW (2005) The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res.* **33** (Database Issue): D418-424.

Alkon DL, Nelson TJ (1990) Specificity of molecular changes in neurons involved in memory storage. *FASEB J.* **4**(6): 1567-1576.

Arenzana-Seisdedos F, Thompson J, Rodriguez MS, Bachelerie F, Thomas D, Hay RT (1995) Inducible nuclear expression of newly synthesized I kappa B alpha negatively regulates DNA-binding and transcriptional activities of NF-kappa B. *Mol Cell Biol.* **15**(5): 2689-2696.

Arenzana-Seisdedos F, Turpin P, Rodriguez M, Thomas D, Hay RT, Virelizier JL, Dargemont C (1997) Nuclear localization of I kappa B alpha promotes active transport of NF-kappa B from the nucleus to the cytoplasm. *J Cell Sci.* **110** (Pt 3): 369-378.

Asensio VC, Campbell IL. (1999) Chemokines in the CNS: plurifunctional mediators in diverse states. *Trends Neurosci.* **22**(11): 504-512.

Bader GD, Betel D, Hogue CW (2003) BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* **31**(1): 248-250.

Bai G, Kusiak JW (1995) Functional analysis of the proximal 5'-flanking region of the N-methyl-D-aspartate receptor subunit gene, NMDAR1. *J Biol Chem.* **270**(13): 7737-7744.

Bajic VB, Chong A, Seah SH, Brusic V (2002a) Intelligent system for vertebrate promoter recognition. *IEEE Intelligent Systems.* **17**(4): 64-70.

Bajic VB, Seah SH, Chong A, Zhang G, Koh JL, Brusic V (2002b) Dragon Promoter Finder: recognition of vertebrate RNA polymerase II promoters. *Bioinformatics.* **18**(1):198-199.

Bajic VB, Seah SH (2003) Dragon gene start finder: an advanced system for finding approximate locations of the start of gene transcriptional units. *Genome Res.* **13**(8): 1923-1929.

Bajic VB, Tan SL, Chong A, Tang S, Strom A, Gustafsson JA, Lin CY, Liu ET (2003) Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes. *Nucleic Acids Res.* **31**(13): 3605-3607.

Bajic VB, Choudhary V**,** Hock CK (2004) Content analysis of the core promoter region of human genes, *In Silico Biology.* **4**:109-125.

Baldini G, Hohl T, Lin HY, Lodish HF (1992) Cloning of a Rab3 isotype predominantly expressed in adipocytes. *Proc Natl Acad Sci U S A.* **89**(11): 5049-5052.

Bal-Price A, Brown GC (2001) Inflammatory neurodegeneration mediated by nitric oxide from activated glia-inhibiting neuronal respiration, causing glutamate release and excitotoxicity. *J Neurosci.* **21**(17): 6480-6491.

Baranes D, Lederfein D, Huang YY, Chen M, Bailey CH, Kandel ER (1998) Tissue plasminogen activator contributes to the late phase of LTP and to synaptic growth in the hippocampal mossy fiber pathway. *Neuron.* **21**(4): 813-825.

Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R (2005) NCBI GEO: mining millions of expression profiles - database and tools. *Nucleic Acids Res.* **33**: D562-566.

Baude A, Nusser Z, Molnar E, McIlhinney RA, Somogyi P (1995) High-resolution immunogold localization of AMPA type glutamate receptor subunits at synaptic and non-synaptic sites in rat hippocampus. *Neuroscience.* **69**(4): 1031-1055.

Beauvais DM, Rapraeger AC (2004) Syndecans in tumor cell adhesion and signaling. *Reprod Biol Endocrinol.* **2**(1): 3.

Bell SD, Kosa PL, Sigler PB, Jackson SP (1999) Orientation of the transcription preinitiation complex in archaea. *Proc Natl Acad Sci U S A.* **96**(24): 13662-13667.

Bennett JA, Dingledine R (1995) Topology profile for a glutamate receptor: three transmembrane domains and a channel-lining reentrant membrane loop. *Neuron.* **14**(2): 373-384.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. *Nucleic Acids Res.* **33**(Database Issue): D34-38.

Benveniste EN (1998) Cytokine actions in the central nervous system. *Cytokine Growth Factor Rev.* **9**(3-4): 259-275.

Bermingham NA, Rauf S, Katsanis N, Martin JE, Hunter AJ, Fisher EM (1998) The immunophilin FKBP4 (FKBP52/FKBP59) maps to the distal short arm of human chromosome 12. *Mamm Genome.* **9**(3): 268.

Bernard V, Somogyi P, Bolam JP (1997) Cellular, subcellular, and subsynaptic distribution of AMPA-type glutamate receptor subunits in the neostriatum of the rat. *J Neurosci.* **17**(2): 819-833.

Bernard V, Bolam JP (1998) Subcellular and subsynaptic distribution of the NR1 subunit of the NMDA receptor in the neostriatum and globus pallidus of the rat: co-localization at synapses with the GluR2/3 subunit of the AMPA receptor. *Eur J Neurosci.* **10**(12): 3721-3736.

Bessis A, Salmon AM, Zoli M, Le Novere N, Picciotto M, Changeux JP (1995) Promoter elements conferring neuron-specific expression of the beta 2-subunit of the neuronal nicotinic acetylcholine receptor studied in vitro and in transgenic mice. *Neuroscience.* **69**(3): 807-819.

Birbach A, Gold P, Binder BR, Hofer E, de Martin R, Schmid JA (2002) Signaling molecules of the NF-kappa B pathway shuttle constitutively between cytoplasm and nucleus. *J Biol Chem.* **277**(13): 10842-10851.

Bliss TV, Gardner-Medwin AR (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the unanaestetized rabbit following stimulation of the perforant path. *J Physiol.* **232**(2): 357-374.

Bliss TV, Lomo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol.* **232**(2): 331-356.

Boehlk S, Fessele S, Mojaat A, Miyamoto NG, Werner T, Nelson EL, Schlondorff D, Nelson PJ (2000) ATF and Jun transcription factors, acting through an Ets/CRE promoter module, mediate lipopolysaccharide inducibility of the chemokine RANTES in monocytic Mono Mac 6 cells. *Eur J Immunol.* **30**(4): 1102-1112.

Borges K, Dingledine R (2001) Functional organization of the GluR1 glutamate receptor promoter. *J Biol Chem.* **276**(28): 25929-25938.

Boulter J, Hollmann M, O'Shea-Greenfield A, Hartley M, Deneris E, Maron C, Heinemann S (1990) Molecular cloning and functional expression of glutamate receptor subunit genes. *Science.* **249**(4972): 1033-1037.

Bourke E, Kennedy EJ, Moynagh PN (2000) Loss of Ikappa B-beta is associated with prolonged NF-kappa B activity in human glial cells. *J Biol Chem.* **275**(51): 39996-40002.

Brené S, Messer C, Nestler EJ (1998) Expression of messenger RNAs encoding ionotropic glutamate receptors in rat brain: regulation by haloperidol. *Neuroscience.* **84**(3): 813-823.

Brené S, Messer C, Okado H, Hartley M, Heinemann SF, Nestler EJ (2000) Regulation of GluR2 promoter activity by neurotrophic factors via a neuron-restrictive silencer element. *Eur J Neurosci.* **12**(5): 1525-1533.

Broutman G, Baudry M (2001) Involvement of the secretory pathway for AMPA receptors in NMDA-induced potentiation in hippocampus. *J Neurosci.* **21**(1): 27-34.

Budde LM, Wu C, Tilman C, Douglas I, Ghosh S (2002) Regulation of IkappaBbeta expression in testis. *Mol Biol Cell.* **13**(12):4179-4194.

Buffa R, Rindi G, Sessa F, Gini A, Capella C, Jahn R, Navone F, De Camilli P, Solcia E (1987) Synaptophysin immunoreactivity and small clear vesicles in neuroendocrine cells and related tumours. *Mol Cell Probes.* **1**(4): 367-381.

Burke TW, Kadonaga JT (1997) The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. *Genes Dev.* **11**(22): 3020-3031.

Burke JR (2003) Targeting I kappa B kinase for the treatment of inflammatory and other disorders. *Curr Opin Drug Discov Devel.* **6**(5): 720-728.

Burnashev N, Khodorova A, Jonas P, Helm PJ, Wisden W, Monyer H, Seeburg PH, Sakmann B (1992a) Calcium-permeable AMPA-kainate receptors in fusiform cerebellar glial cells. *Science.* **256**(5063): 1566-1570.

Burnashev N, Monyer H, Seeburg PH, Sakmann B (1992b) Divalent ion permeability of AMPA receptor channels is dominated by the edited form of a single subunit. *Neuron.* **8**(1): 189-198.

Calakos N, Scheller RH (1994) Vesicle-associated membrane protein and synaptophysin are associated on the synaptic vesicle. *J Biol Chem.* **269**(40): 24534-24537.

Cammarota M, Bevilaqua LR, Bonini JS, Rossatto JI, Medina JH, Izquierdo N (2004) Hippocampal glutamate receptors in fear memory consolidation. *Neurotox Res.* **6**(3): 205-212.

Carcamo J, Buckbinder L, Reinberg D (1991) The initiator directs the assembly of a transcription factor IID-dependent transcription complex. *Proc Natl Acad Sci U S A.* **88**(18): 8052-8056.

Castillo PE, Janz R, Südhof TC, Tzounopoulos T, Malenka RC, Nicoll RA (1997) Rab3A is essential for mossy fibre long-term potentiation in the hippocampus. *Nature.* **388**(6642): 590-593.

Cavin Périer R, Junier T, Bucher P (1998) The Eukaryotic Promoter Database EPD. *Nucleic Acids Res.* **26**(1): 353-357.

Ceccaldi PE, Grohovaz F, Benfenati F, Chieregatti E, Greengard P, Valtorta F (1995) Dephosphorylated synapsin I anchors synaptic vesicles to actin cytoskeleton: an analysis by videomicroscopy. *J Cell Biol.* **128**(5): 905-912.

Chalkley GE, Verrijzer CP (1999) DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J.* **18**(17): 4835-4845.

Chen QK, Hertz GZ, Stormo GD (1995) MATRIX SEARCH 1.0: a computer program that scans DNA sequences for transcriptional elements using a database of weight matrices. *Comput Appl Biosci.* **11**(5): 563-566.

Chen MS, Silverstein AM, Pratt WB, Chinkers M (1996) The tetratricopeptide repeat domain of protein phosphatase 5 mediates binding to glucocorticoid receptor heterocomplexes and acts as a dominant negative mutant. *J Biol Chem.* **271**(50): 32315-32320.

Chen YA, Scheller RH (2001) SNARE-mediated membrane fusion. *Nat Rev Mol Cell Biol.* **2**(2): 98-106.

Chen ZY, Shie JL, Tseng CC (2002) STAT1 is required for IFN-gamma-mediated gut-enriched Kruppel-like factor expression. *Exp Cell Res.* **281**(1): 19-27.

Chew LJ, Yuan X, Scherer SE, Qie L, Huang F, Hayes WP, Gallo V (2001) Characterization of the rat GRIK5 kainate receptor subunit gene promoter and its intragenic regions involved in neural cell specificity. *J Biol Chem.* **276**(45): 42162-42171.

Chi T, Lieberman P, Ellwood K, Carey M (1995) A general mechanism for transcriptional synergy by eukaryotic activators. *Nature.* **377**(6546): 254-257.

Chiao PJ, Miyamoto S, Verma IM (1994) Autoregulation of I kappa B alpha activity. *Proc Natl Acad Sci U S A.* **91**(1): 28-32.

Chin LS, Li L, Greengard P (1994) Neuron-specific expression of the synapsin II gene is directed by a specific core promoter and upstream regulatory elements. *J Biol Chem.* **269**(28): 18507-18513.

Chisamore B, Solc M, Dow K (1996) Excitatory amino acid regulation of astrocyte proteoglycans. *Brain Res Dev Brain Res.* **97**(1): 22-28.

Choi DW (1988) Glutamate neurotoxicity and diseases of the nervous system. *Neuron.* **1**(8): 623-634.

Chong JA, Tapia-Ramirez J, Kim S, Toledo-Aral JJ, Zheng Y, Boutros MC, Altshuller YM, Frohman MA, Kraner SD, Mandel G (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell.* **80**(6): 949-957.

Chong A, Zhang G, Bajic VB (2002) Information and sequence extraction around the 5'-end and translation initiation site of human genes. *In Silico Biol.* **2**(4): 461-465.

Chong A, Zhang G, Bajic VB (2003) FIE2: A program for the extraction of genomic DNA sequences around the start and translation initiation site of human genes. *Nucleic Acids Res.* **31**(13): 3546-3553.

Clarke NP, Bolam JP (1998) Distribution of glutamate receptor subunits at neurochemically characterized synapses in the entopeduncular nucleus and subthalamic nucleus of the rat. *J Comp Neurol.* **397**(3): 403-420.

Coghlan VM, Perrino BA, Howard M, Langeberg LK, Hicks JB, Gallatin WM, Scott JD (1995) Association of protein kinase A and protein phosphatase 2B with a common anchoring protein. *Science.* **267**(5194): 108-111.

Coleman SL, Buckland PR, Hoogendoorn B, Guy C, Smith K, O'Donovan MC (2002) Experimental analysis of the annotation of promoters in the public database. *Hum. Mol. Genet.* **11**(16): 1817-1821.

Collins JE, Goward ME, Cole CG, Smink LJ, Huckle EJ, Knowles S, Bye JM, Beare DM, Dunham I (2003) Reevaluating human gene annotation: a second-generation analysis of chromosome 22. *Genome Res.* **13**(1): 27-36.

Condorelli DF, Dell'Albani P, Amico C, Kaczmarek L, Nicoletti F, Lukasiuk K, Stella AM (1993) Induction of primary response genes by excitatory amino acid receptor agonists in primary astroglial cultures. *J Neurochem.* **60**(3): 877-885.

Cooney RN (2002) Suppressors of cytokine signaling (SOCS): inhibitors of the JAK/STAT pathway. *Shock.* **17**(2): 83-90.

Cramer P (2002) Multisubunit RNA polymerases. *Curr Opin Struct Biol.* **12**(1): 89-97.

Cramer P (2004) RNA polymerase II structure: from core to functional complexes. *Curr Opin Genet Dev.* **14**(2): 218-226.

Curutchet P, Bochet P, Prado de Carvalho L, Lambolez B, Stinnakre J, Rossier J (1992) In the GluR1 glutamate receptor subunit a glutamine to histidine point mutation suppresses inward rectification but not calcium permeability. *Biochem Biophys Res Commun.* **182**(3):1089-1093.

Czar MJ, Owens-Grillo JK, Dittmar KD, Hutchison KA, Zacharek AM, Leach KL, Deibel MR Jr, Pratt WB (1994) Characterization of the protein-protein interactions determining the heat shock protein (hsp90.hsp70.hsp56) heterocomplex. *J Biol Chem.* **269**(15): 11155-11161.

Czar MJ, Galigniana MD, Silverstein AM, Pratt WB (1997) Geldanamycin, a heat shock protein 90-binding benzoquinone ansamycin, inhibits steroid-dependent translocation of the glucocorticoid receptor from the cytoplasm to the nucleus. *Biochemistry.* **36**(25): 7776-7785.

Dabiri GA, Lai F, Drakas RA, Nishikura K (1996) Editing of the GLuR-B ion channel RNA in vitro by recombinant double-stranded RNA adenosine deaminase. *EMBO J.* **15**(1): 34-45.

Darchen F, Zahraoui A, Hammel F, Monteils MP, Tavitian A, Scherman D (1990) Association of the GTP-binding protein Rab3A with bovine adrenal chromaffin granules. *Proc Natl Acad Sci U S A.* **87**(15): 5692-5696.

Daukas G, Zigmond SH (1985) Inhibition of receptor-mediated but not fluid-phase endocytosis in polymorphonuclear leukocytes. *J Cell Biol.* **101**(5 Pt 1): 1673-1679.

Dawson TM, Steiner JP, Dawson VL, Dinerman JL, Uhl GR, Snyder SH (1993) Immunosuppressant FK506 enhances phosphorylation of nitric oxide synthase and protects against glutamate neurotoxicity. *Proc Natl Acad Sci U S A.* **90**(21): 9808-9812.

Dawson TM, Steiner JP, Lyons WE, Fotuhi M, Blue M, Snyder SH (1994) The immunophilins, FK506 binding protein and cyclophilin, are discretely localized in the brain: relationship to calcineurin. *Neuroscience.* **62**(2): 569-580.

DeBello WM, O'Connor V, Dresbach T, Whiteheart SW, Wang SS, Schweizer FE, Betz H, Rothman JE, Augustine GJ (1995) SNAP-mediated protein-protein interactions essential for neurotransmitter release. *Nature.* **373**(6515): 626-630.

Diamond MI, Miner JN, Yoshinaga SK, Yamamoto KR (1990) Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. *Science.* **249**(4974): 1266-1272.

DiDonato JA, Hayakawa M, Rothwarf DM, Zandi E, Karin M (1997) A cytokine-responsive IkappaB kinase that activates the transcription factor NF-kappaB. *Nature.* **388**(6642): 548-554.

Dingledine R, Hume RI, Heinemann SF (1992) Structural determinants of barium permeation and rectification in non-NMDA glutamate receptor channels. *J Neurosci.* **12**(10): 4080-4087.

Donaldson I, Martin J, de Bruijn B, Wolting C, Lay V, Tuekam B, Zhang S, Baskin B, Bader GD, Michalickova K, Pawson T, Hogue CW (2003) PreBIND and Textomy--mining the biomedical literature for protein-protein interactions using a support vector machine. *BMC Bioinformatics.* **4**(1): 11.

Douziech M, Coin F, Chipoulet JM, Arai Y, Ohkuma Y, Egly JM, Coulombe B (2000) Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol Cell Biol.* **20**(21): 8168-8177.

Dowling JE (1987) The retina: an approachable part of the brain. *The Belknap Press of Harvard University Press.*

D'Souza SM, Brown IR (1998) Constitutive expression of heat shock proteins Hsp90, Hsc70, Hsp70 and Hsp60 in neural and non-neural tissues of the rat during postnatal development. *Cell Stress Chaperones.* **3**(3): 188-199.

Durand GM, Kovalchuk Y, Konnerth A (1996) Long-term potentiation and functional synapse induction in developing hippocampus. *Nature.* **381**(6577): 71-75.

Edelmann L, Hanson PI, Chapman ER, Jahn R (1995) Synaptobrevin binding to synaptophysin: a potential mechanism for controlling the exocytotic fusion machine. *EMBO J.* **14**(2): 224-231.

Elkon R, Zeller KI, Linhart C, Dang CV, Shamir R, Shiloh Y. (2004) In silico identification of transcriptional regulators associated with c-Myc. *Nucleic Acids Res.* **32**(17):4955-61.

Endo TA, Masuhara M, Yokouchi M, Suzuki R, Sakamoto H, Mitsui K, Matsumoto A, Tanimura S, Ohtsubo M, Misawa H, Miyazaki T, Leonor N, Taniguchi T, Fujita T, Kanakura Y, Komiya S, Yoshimura A (1997) A new protein containing an SH2 domain that inhibits JAK kinases. *Nature*. **387**(6636): 921-924.

Engert F, Bonhoeffer T (1999) Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature.* **399**(6731): 66-70.

Eshhar N, Petralia RS, Winters CA, Niedzielski AS, Wenthold RJ (1993) The segregation and expression of glutamate receptor subunits in cultured hippocampal neurons. *Neuroscience.* **57**(4): 943-964.

Fairley JA, Evans R, Hawkes NA, Roberts SG (2002) Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection. *Mol Cell Biol.* **22**(19): 6697-6705.

Feaver WJ, Gileadi O, Li Y, Kornberg RD (1991) CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell.* **67**(6): 1223-1230.

Firulli AB, Olson EN (1997) Modular regulation of muscle gene transcription: a mechanism for muscle cell diversity. *Trends Genet.* **13**(9): 364-369.

Fischer von Mollard G, Mignery GA, Baumert M, Perin MS, Hanson TJ, Burger PM, Jahn R, Südhof TC (1990) rab3 is a small GTP-binding protein exclusively localized to synaptic vesicles. *Proc Natl Acad Sci U S A.* **87**(5): 1988-1992.

Frey U, Muller M, Kuhl D (1996) A different form of long-lasting potentiation revealed in tissue plasminogen activator mutant mice. *J Neurosci.* **16**(6): 2057-2063.

Friedman LK, Pellegrini-Giampietro DE, Sperber EF, Bennett MV, Moshe SL, Zukin RS (1994) Kainate-induced status epilepticus alters glutamate and GABAA receptor gene expression in adult rat hippocampus: an in situ hybridization study. *J Neurosci.* **14**(5 Pt 1): 2697-2707.

Fykse EM, Takei K, Walch-Solimena C, Geppert M, Jahn R, De Camilli P, Südhof TC (1993) Relative properties and localizations of synaptic vesicle protein isoforms: the case of the synaptophysins. *J Neurosci.* **13**(11): 4997-5007.

Galas D and Schmitz A. (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5**(9):3157-70.

Galigniana MD, Radanyi C, Renoir JM, Housley PR, Pratt WB (2001) Evidence that the peptidylprolyl isomerase domain of the hsp90-binding immunophilin FKBP52 is involved in both dynein interaction and glucocorticoid receptor movement to the nucleus. *J Biol Chem.* **276**(18): 14884-14889.

Galli T, McPherson PS, De Camilli P (1996) The V0 sector of the V-ATPase, synaptobrevin, and synaptophysin are associated on synaptic vesicles in a Triton X-100-resistant, freeze-thawing sensitive, complex. *J Biol Chem.* **271**(4): 2193-2198.

Ganchi PA, Sun SC, Greene WC, Ballard DW (1992) I kappa B/MAD-3 masks the nuclear localization signal of NF-kappa B p65 and requires the transactivation domain to inhibit NF-kappa B p65 DNA binding. *Mol Biol Cell.* **3**(12): 1339-1352.

Garthwaite J, Charles SL, Chess-Williams R (1988) Endothelium-derived relaxing factor release on activation of NMDA receptors suggests role as intercellular messenger in the brain. *Nature.* **336**(6197): 385-388.

Gaur U, Aggarwal BB (2003) Regulation of proliferation, survival and apoptosis by members of the TNF superfamily. *Biochem Pharmacol.* **66**(8): 1403-1408.

Geiger JH, Hahn S, Lee S, Sigler PB (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science.* **272**(5263): 830-836.

Geppert M, Bolshakov VY, Siegelbaum SA, Takei K, De Camilli P, Hammer RE, Südhof TC (1994) The role of Rab3A in neurotransmitter release. *Nature.* **369**(6480): 493-497.

Geppert M, Goda Y, Stevens CF, Südhof TC (1997) The small GTP-binding protein Rab3A regulates a late step in synaptic vesicle fusion. *Nature.* **387**(6635): 810-814.

Gerges NZ, Tran IC, Backos DS, Harrell JM, Chinkers M, Pratt WB, Esteban JA (2004) Independent functions of hsp90 in neurotransmitter release and in the continuous synaptic cycling of AMPA receptors. *J Neurosci.* **24**(20): 4758-4766.

Gerst JE (2003) SNARE regulators: matchmakers and matchbreakers. *Biochim Biophys Acta.* **1641**(2-3): 99-110.

Gleeson PA, Lock JG, Luke MR, Stow JL (2004) Domains of the TGN: Coats, Tethers and G Proteins. *Traffic.* **5**(5): 315-326.

Gollub J, Ball CA, Binkley G, Demeter J, Finkelstein DB, Hebert JM, Hernandez-Boussard T, Jin H, Kaloper M, Matese JC, Schroeder M, Brown PO, Botstein D, Sherlock G (2003) The Stanford Microarray Database: data access and quality assessment tools. *Nucleic Acids Res* **31**(1): 94-96.

Gorter JA, Petrozzino JJ, Aronica EM, Rosenbaum DM, Opitz T, Bennett MV, Connor JA, Zukin RS (1997) Global ischemia induces downregulation of Glur2 mRNA and increases AMPA receptor-mediated Ca^{2+} influx in hippocampal CA1 neurons of gerbil. *J Neurosci.* **17**(16): 6179-6188.

Green MR (2000) TBP-associated factors (TAFIIs): multiple, selective transcriptional mediators in common complexes. *Trends Biochem Sci.* **25**(2):59-63.

Greengard P, Jen J, Nairn AC, Stevens CF (1991) Enhancement of the glutamate response by cAMP-dependent protein kinase in hippocampal neurons. *Science.* **253**(5024): 1135-1138.

Gregersen R, Lambertsen K, Finsen B (2000) Microglia and macrophages are the major source of tumor necrosis factor in permanent middle cerebral artery occlusion in mice. *J Cereb Blood Flow Metab.* **20**(1): 53-65.

Grenningloh G, Gundelfinger E, Schmitt B, Betz H, Darlison MG, Barnard EA, Schofield PR, Seeburg PH (1987) Glycine vs GABA receptors. *Nature.* **330**(6143): 25-26.

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of Arabidopsis transcription factors. *Bioinformatics.* [Epub ahead of print].

Ha I, Roberts S, Maldonado E, Sun X, Kim LU, Green M, Reinberg D (1993) Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes Dev.* **7**(6): 1021-1032.

Halees AS, Leyfer D, Weng Z (2003) PromoSer: A large-scale mammalian promoter and transcription start site identification service. *Nucleic Acids Res.* **31**(13): 3554-3559.

Hampsey M (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol Mol Biol Rev.* **62**(2): 465-503.

Hanson PI, Roth R, Morisaki H, Jahn R, Heuser JE (1997) Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. *Cell.* **90**(3): 523-535.

Harada K, Masuda S, Hirano M, Nakanuma Y (2003) Reduced expression of syndecan-1 correlates with histologic dedifferentiation, lymph node metastasis, and poor prognosis in intrahepatic cholangiocarcinoma. *Hum Pathol.* **34**(9): 857-863.

Hay JC, Scheller RH (1997) SNAREs and NSF in targeted membrane fusion. *Curr Opin Cell Biol.* **9**(4): 505-512.

Hayashi T, Umemori H, Mishina M, Yamamoto T (1999) The AMPA receptor interacts with and signals through the protein tyrosine kinase Lyn. *Nature.* **397**(6714): 72-76.

Hayashi Y, Shi SH, Esteban JA, Piccini A, Poncer JC, Malinow R (2000) Driving AMPA receptors into synapses by LTP and CaMKII: requirement for GluR1 and PDZ domain interaction. *Science.* **287**(5461): 2262-2267.

Heinemeyer T, Chen X, Karas H, Kel AE, Kel OV, Liebich I, Meinhardt T, Reuter I, Schacherer F, Wingender E (1999) Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. *Nucleic Acids Res.* **27**(1): 318-322.

Heuser JE, Anderson RG (1989) Hypertonic media inhibit receptor-mediated endocytosis by blocking clathrin-coated pit formation. *J Cell Biol.* **108**(2):389-400.

Hintsch G, Zurlinden A, Meskenaite V, Steuble M, Fink-Widmer K, Kinter J, Sonderegger P (2002) The calsyntenins--a family of postsynaptic membrane proteins with distinct neuronal expression patterns. *Mol Cell Neurosci.* **21**(3): 393-409.

Hirata A, Yoshida S, Inoue N, Matsumoto-Miyai K, Ninomiya A, Taniguchi M, Matsuyama T, Kato K, Iizasa H, Kataoka Y, Yoshida N, Shiosaka S (2001) Abnormalities of synapses and neurons in the hippocampus of neuropsin-deficient mice. *Mol Cell Neurosci.* **17**(3): 600-610.

Hirsch DB, Steiner JP, Dawson TM, Mammen A., Hayek E. and Snyder SH (1993) Neurotransmitter Release Regulated by Nitric Oxide in PC-12 Cells and Brain Synaptosomes. *Curr.Biol.* **3**: 749-754.

Hohnke CD, Oray S, Sur M (2000) Activity-dependent patterning of retinogeniculate axons proceeds with a constant contribution from AMPA and NMDA receptors. *J Neurosci.* **20**(21): 8051-8060..

Hollmann M, O'Shea-Greenfield A, Rogers SW, Heinemann S (1989) Cloning by functional expression of a member of the glutamate receptor family. *Nature.* **342**(6250): 643-648.

Hollmann M, Hartley M, Heinemann S (1991) Ca^{2+} permeability of KA-AMPA--gated glutamate receptor channels depends on subunit composition. *Science.* **252**(5007): 851-853.

Hollmann M, Maron C, Heinemann S (1994) N-glycosylation site tagging suggests a three transmembrane domain topology for the glutamate receptor GluR1. *Neuron.* **13**(6): 1331-1343.

Hollmann M, Heinemann S (1994) Cloned glutamate receptors. *Annu Rev Neurosci.* **17**: 31-108.

Hope BT (1998) Cocaine and the AP-1 transcription factor complex. *Ann N Y Acad Sci.* **844**: 1-6.

Horikawa HP, Kneussel M, El Far O, Betz H (2002) Interaction of synaptophysin with the AP-1 adaptor protein gamma-adaptin. *Mol Cell Neurosci.* **21**(3): 454-462.

Huang YY, Bach ME, Lipp HP, Zhuo M, Wolfer DP, Hawkins RD, Schoonjans L, Kandel ER, Godfraind JM, Mulligan R, Collen D, Carmeliet P (1996) Mice lacking the gene encoding tissue-type plasminogen activator show a selective interference with late-phase long-term potentiation in both Schaffer collateral and mossy fiber pathways. *Proc Natl Acad Sci U S A.* **93**(16): 8699-8704.

Huang Y, Myers SJ, Dingledine R (1999) Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nat Neurosci.* **2**(10): 867-872.

Huang TT, Miyamoto S (2001) Postrepression activation of NF-kappaB requires the amino-terminal nuclear export signal specific to IkappaBalpha. *Mol Cell Biol.* **21**(14): 4737-4747.

Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, Cox T, Cuff J, Curwen V, Down T, Durbin R, Eyras E, Gilbert J, Hammond M, Huminiecki L, Kasprzyk A, Lehvaslaiho H, Lijnzaad P, Melsopp C, Mongin E, Pettett R, Pocock M, Potter S, Rust A, Schmidt E, Searle S, Slater G, Smith J, Spooner W, Stabenau A, Stalker J, Stupka E, Ureta-Vidal A, Vastrik I, Clamp M (2002) The Ensembl genome database project. *Nucleic Acids Res.* **30**(1): 38-41.

Hume RI, Dingledine R, Heinemann SF (1991) Identification of a site in glutamate receptor subunits that controls calcium permeability. *Science.* **253**(5023): 1028-1031.

Inatani M, Irie F, Plump AS, Tessier-Lavigne M, Yamaguchi, Y (2003) Mammalian brain morphogenesis amd midline axon guidance require heparan sulfate. *Science.* **302**(5647): 1044-1046.

Inglis FM, Crockett R, Korada S, Abraham WC, Hollmann M, Kalb RG (2002) The AMPA receptor subunit GluR1 regulates dendritic architecture of motor neurons. *J Neurosci.* **22**(18): 8042-8051.

Isaac JTR, Nicoll RA, Malenka RC (1995) Evidence for silent synapses: implications for the expression of LTP. *Neuron.* **15**: 427-434.

Jahn R, Schiebler W, Ouimet C, Greengard P (1985) A 38,000-dalton membrane protein (p38) present in synaptic vesicles. *Proc Natl Acad Sci U S A.* **82**(12): 4137-4141.

Jakowec MW, Fox AJ, Martin LJ, Kalb RG (1995a) Quantitative and qualitative changes in AMPA receptor expression during spinal cord development. *Neuroscience.* **67**(4): 893-907.

Jakowec MW, Yen L, Kalb RG (1995b) In situ hybridization analysis of AMPA receptor subunit gene expression in the developing rat spinal cord. *Neuroscience.* **67**(4): 909-920.

Jensen FE (2002) The role of glutamate receptor maturation in perinatal seizures and brain injury. *Int J Dev Neurosci.* **20**(3-5): 339-347.

Jiang Y, Chen D, Lyu SC, Ling X, Krensky AM, Clayberger C (2002) DQ 65-79, a peptide derived from HLA class II, induces I kappa B expression. *J Immunol.* **168**(7): 3323-3328.

Johnson DR, Douglas I, Jahnke A, Ghosh S, Pober JS (1996) A sustained reduction in IkappaB-beta may contribute to persistent NF-kappaB activation in human endothelial cells. *J Biol Chem.* **271**(27): 16317-16322.

Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science.* **316**: 1497-1502.

Jonas P, Racca C, Sakmann B, Seeburg PH, Monyer H (1994) Differences in Ca^{2+} permeability of AMPA-type glutamate receptor channels in neocortical neurons caused by differential GluR-B subunit expression. *Neuron.* **12**(6): 1281-1289.

Jothi R, Cuddapah S, Barski A, Cui K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.* **36**: 5221-5231.

Kalb RG, Lidow MS, Halsted MJ, Hockfield S (1992) N-methyl-D-aspartate receptors are transiently expressed in the developing spinal cord ventral horn. *Proc Natl Acad Sci U S A.* **89**(18): 8502-8506.

Kamboj SK, Swanson GT, Cull-Candy SG (1995) Intracellular spermine confers rectification on rat calcium-permeable AMPA and kainate receptors. *J Physiol.* **486** (Pt 2): 297-303.

Kandel ER (1997) Genes, synapses, and long-term memory. *J Cell Physiol.* **173**(2): 124-125.

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1): 27-30.

Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**(Database issue): D277-280.

Kaneider NC, Dunzendorfer S, Wiedermann CJ (2004) Heparan sulfate proteoglycans are involved in opiate receptor-mediated cell migration. *Biochemistry.* **43**(1): 237-244.

Kanner BI, Sharon I (1978) Active transport of L-glutamate by membrane vesicles isolated from rat brain. *Biochemistry.* **17**(19): 3949-3953.

Kanz C, Aldebert P, Althorpe N, Baker W, Baldwin A, Bates K, Browne P, van den Broek A, Castro M, Cochrane G, Duggan K, Eberhardt R, Faruque N, Gamble J, Diez FG, Harte N, Kulikova T, Lin Q, Lombard V, Lopez R, Mancuso R, McHale M, Nardone F, Silventoinen V, Sobhany S, Stoehr P, Tuli MA, Tzouvara K, Vaughan R, Wu D, Zhu W, Apweiler R (2005) The EMBL Nucleotide Sequence Database. *Nucleic Acids Res.* **33** (Database Issue): D29-33.

Karin M, Ben-Neriah Y (2000) Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu Rev Immunol.* **18**: 621-663.

Katz LC, Constantine-Paton M (1988) Relationships between segregated afferents and postsynaptic neurones in the optic tectum of three-eyed frogs. *J Neurosci.* **8**(9): 3160-3180.

Kazlauskas A, Poellinger L, Pongratz I (2000) The immunophilin-like protein XAP2 regulates ubiquitination and subcellular localization of the dioxin receptor. *J Biol Chem.* **275**(52): 41317-41324.

Kazlauskas A, Sundstrom S, Poellinger L, Pongratz I (2001) The hsp90 chaperone complex regulates intracellular localization of the dioxin receptor. *Mol Cell Biol.* **21**(7): 2594-2607.

Keinänen K, Wisden W, Sommer B, Werner P, Herb A, Verdoorn TA, Sakmann B, Seeburg PH (1990) A family of AMPA-selective glutamate receptors. *Science.* **249**(4968): 556-560.

Kel OV, Romaschenko AG, Kel AE, Wingender E, Kolchanov NA (1995) A compilation of composite regulatory elements affecting gene transcription in vertebrates. *Nucleic Acids Res.* **23**(20): 4097-4103.

Kel A, Kel-Margoulis O, Babenko V, Wingender E (1999) Recognition of NFATp/AP-1 composite elements within genes induced upon the activation of immune cells. *J Mol Biol.* **288**(3): 353-376.

Kel-Margoulis OV, Romashchenko AG, Kolchanov NA, Wingender E, Kel AE (2000) COMPEL: a database on composite regulatory elements providing combinatorial transcriptional regulation. *Nucleic Acids Res.* **28**(1): 311-315.

Kel-Margoulis OV, Kel AE, Reuter I, Deineko IV, Wingender E (2002) TRANSCompel: a database on composite regulatory elements in eukaryotic genes. *Nucleic Acids Res.* **30**(1): 332-334.

Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **31**(13): 3576-3579.

Keller BU, Hollmann M, Heinemann S, Konnerth A (1992) Calcium influx through subunits GluR1/GluR3 of kainate/AMPA receptor channels is regulated by cAMP dependent protein kinase. *EMBO J.* **11**(3): 891-896.

Kelz MB, Chen J, Carlezon WA Jr, Whisler K, Gilden L, Beckmann AM, Steffen C, Zhang YJ, Marotti L, Self DW, Tkatch T, Baranauskas G, Surmeier DJ, Neve RL, Duman RS, Picciotto MR, Nestler EJ (1999) Expression of the transcription factor deltaFosB in the brain controls sensitivity to cocaine. *Nature.* **401**(6750): 272-276.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. *Genome Res.* **12**(6): 996-1006.

Kent WJ, Hsu F, Karolchik D, Kuhn RM, Clawson H, Trumbower H, Haussler D (2005) Exploring relationships and mining data with the UCSC Gene Sorter. *Genome Res.* **15**(5): 737-741.

Kessler JP, Baude A (1999) Distribution of AMPA receptor subunits GluR1-4 in the dorsal vagal complex of the rat: a light and electron microscope immunocytochemical study. *Synapse.* **34**(1):55-67.

Kim U, Nishikura K (1993) Double-stranded RNA adenosine deaminase as a potential mammalian RNA editing factor. *Semin Cell Biol.* **4**(4): 285-293.

Kim Y, Geiger JH, Hahn S, Sigler PB (1993a) Crystal structure of a yeast TBP/TATA-box complex. *Nature.* **365**(6446): 512-520.

Kim JL, Nikolov DB, Burley SK (1993b) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature.* **365**(6446): 520-527.

Kim TK, Ebright RH, Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science.* **288**(5470): 1418-1422.

Kim J, Seo J, Lee YS, Kim S (2005) TFExplorer: integrated analysis database for predicted transcription regulatory elements. *Bioinformatics.* **21**(4): 548-550.

Kiyama H, Sato K, Tohyama M (1993) Characteristic localization of non-NMDA type glutamate receptor subunits in the rat pituitary gland. *Brain Res Mol Brain Res.* **19**(3): 262-268.

Klein M, Pieri I, Uhlmann F, Pfizenmaier K, Eisel U (1998) Cloning and characterization of promoter and 5'-UTR of the NMDA receptor subunit epsilon 2: evidence for alternative splicing of 5'-non-coding exon. *Gene.* **208**(2): 259-269.

Knuppel R, Dietze P, Lehnberg W, Frech K, Wingender E (1994) TRANSFAC retrieval program: a network model database of eukaryotic transcription regulating sequences and proteins. *J Comput Biol.* **1**(3): 191-198.

Koh DS, Burnashev N, Jonas P (1995) Block of native Ca(2+)-permeable AMPA receptors in rat brain by intracellular polyamines generates double rectification. *J Physiol.* **486** (Pt 2): 305-312.

Köhler M, Kornau HC, Seeburg PH (1994) The organization of the gene for the functionally dominant alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor subunit GluR-B. *J Biol Chem.* **269**(26): 17367-17370.

Kokubo T, Swanson MJ, Nishikawa JI, Hinnebusch AG, Nakatani Y (1998) The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol Cell Biol.* **18**(2):1003-1012.

Krebs DL, Hilton DJ (2000) SOCS: physiological suppressors of cytokine signaling. *J Cell Sci.* **113** (Pt 16): 2813-2819.

Kucharczak J, Simmons MJ, Fan Y, Gelinas C (2003) To be, or not to be: NF-kappaB is the answer--role of Rel/NF-kappaB in the regulation of apoptosis. *Oncogene.* **22**(56): 8961-8982.

Ladner KJ, Caligiuri MA, Guttridge DC (2003) Tumor necrosis factor-regulated biphasic activation of NF-kappa B is required for cytokine-induced loss of skeletal muscle gene products. *J Biol Chem.* **278**(4): 2294-2303.

Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12**(1): 34-44.

Lam E, Martin M, Wiederrecht G (1995) Isolation of a cDNA encoding a novel human FK506-binding protein homolog containing leucine zipper and tetratricopeptide repeat motifs. *Gene.* **160**(2): 297-302.

Lax P, Limatola C, Fucile S, Trettel F, Di Bartolomeo S, Renzi M, Ragozzino D, Eusebi F (2002) Chemokine receptor CXCR2 regulates the functional properties of AMPA-type glutamate receptor GluR1 in HEK cells. *J Neuroimmunol.* **129**(1-2): 66-73.

Lee Y, Allport V, Sykes A, Lindstrom T, Slater D, Bennett P (2003) The effects of labour and of interleukin 1 beta upon the expression of nuclear factor kappa B related proteins in human amnion. *Mol Hum Reprod.* **9**(4): 213-218.

Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P (2004) Coexpression analysis of human genes across many microarray data sets *Genome Res.* **14**(6): 1085-1094.

Lefstin JA, Yamamoto KR (1998) Allosteric effects of DNA on transcriptional regulators. *Nature.* **392**(6679): 885-888.

Lennon G, Auffray C, Polymeropoulos M, Soares MB (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics.* **33**(1): 151-152.

Lenzen CU, Steinmann D, Whiteheart SW, Weis WI (1998) Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell.* **94**(4): 525-536.

Li Q, Verma IM (2002) NF-kappaB regulation in the immune system. *Nat Rev Immunol.* **2**(10): 725-734.

Liao D-Z, Hessler NA, Malinow R (1995) Activation of postsynaptic silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature.* **375**: 400-404.

Liao D, Zhang X, O'Brien R, Ehlers MD, Huganir RL (1999) Regulation of morphological postsynaptic silent synapses in developing hippocampal neurons. *Nat Neurosci.* **2**(1): 37-43.

Liao D, Scannevin RH, Huganir R (2001) Activation of silent synapses by rapid activity-dependent synaptic recruitment of AMPA receptors. *J Neurosci.* **21**(16): 6008-6017.

Liman ER, Knapp AG, Dowling JE (1989) Enhancement of kainate-gated currents in retinal horizontal cells by cyclic AMP-dependent protein kinase. *Brain Res.* **481**(2): 399-402.

Limatola C, Ciotti MT, Mercanti D, Santoni A, Eusebi F (2002) Signaling pathways activated by chemokine receptor CXCR2 and AMPA-type glutamate receptors and involvement in granule cells survival. *J Neuroimmunol.* **123**(1-2): 9-17.

Limatola C, Di Bartolomeo S, Trettel F, Lauro C, Ciotti MT, Mercanti D, Castellani L, Eusebi F (2003) Expression of AMPA-type glutamate receptors in HEK cells and cerebellar granule neurons impairs CXCL2-mediated chemotaxis. *J Neuroimmunol.* **134**(1-2): 61-71.

Lin CG, Lin YC, Liu HW, Kao LS (1997) Characterization of Rab3A, Rab3B and Rab3C: different biochemical properties and intracellular localization in bovine chromaffin cells. *Biochem J.* **324** (Pt 1): 85-90.

Lin, C.Y., Ström, A., Vega, V.B., Kong, S.L., Yeo, A.L., Thomsen, J.S., Chan, W.C., Doray, B., Bangarusamy, D.K., Ramasamy, A., Vergara, L.A., Tang, S., Chong, A., Bajic, V.B., Miller, L.D., Gustafsson, J.A., Liu, E.T. (2004) Discovery of estrogen receptor alpha receptor alpha target genes and response elements in breast tumor cells. *Genome Biol.* **5**(9): R66.

Lin RC, Scheller RH (1997) Structural organization of the synaptic exocytosis core complex. *Neuron.* **19**(5): 1087-1094.

Lin RC, Scheller RH (2000) Mechanisms of synaptic vesicle exocytosis. *Annu Rev Cell Dev Biol.* **16**: 19-49.

Littlefield O, Korkhin Y, Sigler PB (1999) The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc Natl Acad Sci U S A.* **96**(24): 13668-13673.

Liu J, Farmer JD Jr, Lane WS, Friedman J, Weissman I, Schreiber SL (1991) Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell.* **66**(4): 807-815.

Liu T, Clark RK, McDonnell PC, Young PR, White RF, Barone FC, Feuerstein GZ (1994) Tumor necrosis factor-alpha expression in ischemic neurons. *Stroke.* **25**(7): 1481-1488.

Liu J, Shworak NW, Fritze LM, Edelberg JM, Rosenberg RD (1996) Purification of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *J Biol Chem.* **271**(43): 27072-27082.

Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, Kay LE, Nakatani Y, Ikura M (1998) Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell.* **94**(5): 573-583.

Liu J, Shworak NW, Sinay P, Schwartz JJ, Zhang L, Fritze LM, Rosenberg RD (1999a) Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities. *J Biol Chem.* **274**(8): 5185-5192.

Liu J, Shriver Z, Blaiklock P, Yoshida K, Sasisekharan R, Rosenberg RD (1999b) Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3A sulfates N-unsubstituted glucosamine residues. *J Biol Chem.* **274**(53): 38155-38162.

Liu M, Mendicino M, Ning Q, Ghanekar A, He W, McGilvray I, Shalev I, Pivato D, Clark DA, Phillips MJ, Levy GA (2006) Cytokine-induced hepatic apoptosis is dependent on FGL2/fibroleukin: the role of Sp1/Sp3 and STAT1/PU.1 composite cis elements. *J Immunol.* **176**(11): 7028-7038.

Lledo PM, Vernier P, Vincent JD, Mason WT, Zorec R (1993) Inhibition of Rab3B expression attenuates Ca(2+)-dependent exocytosis in rat anterior pituitary cells. *Nature.* **364**(6437): 540-544.

Lledo PM, Zhang X, Südhof TC, Malenka RC, Nicoll RA (1998) Postsynaptic membrane fusion and long-term potentiation. *Science.* **279**(5349): 399-403.

Llinás R, Gruner JA, Sugimori M, McGuinness TL, Greengard P (1991) Regulation by synapsin I and Ca(2+)-calmodulin-dependent protein kinase II of the transmitter release in squid giant synapse. *J Physiol.* **436**: 257-282.

Lomeli H, Mosbacher J, Melcher T, Hoger T, Geiger JR, Kuner T, Monyer H, Higuchi M, Bach A, Seeburg PH (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science.* **266**(5191): 1709-1713.

Loots GG, Ovcharenko I (2004) rVISTA 2.0: evolutionary analysis of transcription factor binding sites. *Nucleic Acids Res.* **32**(Web Server issue): W217-221.

Lu H, Zawel L, Fisher L, Egly JM, Reinberg D (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature.* **358**(6388): 641-645.

Lu L, Lu H, Skolnick J (2002) MULTIPROSPECTOR: an algorithm for the prediction of protein-protein interactions by multimeric threading. *Proteins.* **49**(3): 350-364.
Lu L, Arakaki AK, Lu H, Skolnick J (2003) Multimeric threading-based prediction of protein-protein interactions on a genomic scale: application to the Saccharomyces cerevisiae proteome. *Genome Res.* **13**(6A): 1146-1154.

Maglott, D.R., Katz, K.S., Sicotte, H., Pruitt, K.D. (2000) NCBI's LocusLink and RefSeq. *Nucleic Acids Res.,* **28**(1): 126-128.

Malek S, Chen Y, Huxford T, Ghosh G (2001) IkappaBbeta, but not IkappaBalpha, functions as a classical cytoplasmic inhibitor of NF-kappaB dimers by masking both NF-kappaB nuclear localization sequences in resting cells. *J Biol Chem.* **276**(48): 45225-45235.

Maletic-Savatic M, Koothan T, Malinow R (1998) Calcium-evoked dendritic exocytosis in cultured hippocampal neurons. Part II: mediation by calcium/calmodulin-dependent protein kinase II. *J Neurosci.* **18**(17): 6814-6821.

Marc RE, Lam DM (1981) Uptake of aspartic and glutamic acid by photoreceptors in goldfish retina. *Proc Natl Acad Sci U S A.* **78**(11): 7185-7189.

Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D (1999) Detecting protein function and protein-protein interactions from genome sequences. *Science.* **285**(5428): 751-753.

Marcotte EM, Xenarios I, Eisenberg D (2001) Mining literature for protein-protein interactions. *Bioinformatics.* **17**(4): 359-363.

Marcoux FW, Weber ML, Probert AW Jr, Dominick MA (1992) Hypoxic neurodegeneration in culture: calcium influx, electron microscopy, and neuroprotection with excitatory amino acid antagonists. *Ann N Y Acad Sci.* **648**: 303-305.

Maren S, Tocco G, Standley S, Baudry M, Thompson RF (1993) Postsynaptic factors in the expression of long-term potentiation (LTP): increased glutamate receptor binding following LTP induction in vivo. *Proc Natl Acad Sci U S A.* **90**(20): 9654-9658.

Marqueze-Pouey B, Wisden W, Malosio ML, Betz H (1991) Differential expression of synaptophysin and synaptoporin mRNAs in the postnatal rat central nervous system. *J Neurosci.* **11**(11): 3388-3397.

Martin JB, Beal MF (1992) Huntington's disease and neurotoxins. *Ann N Y Acad Sci.* **648**: 169-175.

Martin LJ, Blackstone CD, Levey AI, Huganir RL, Price DL (1993) AMPA glutamate receptor subunits are differentially distributed in rat brain. *Neuroscience.* **53**(2): 327-358.

Maruyama K, Sugano S (1994) Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene.* **138**: 171-174.

Mascareno E, Dhar M, Siddiqui MA (1998) Signal transduction and activator of transcription (STAT) protein-dependent activation of angiotensinogen promoter: a cellular signal for hypertrophy in cardiac muscle. *Proc Natl Acad Sci U S A.* **95**(10): 5590-5594.

Massey SC (1990) Cell types using glutamate as a neurotransmitter in the vertebrate retina. In: Progress in retinal research. (N.N. Osborne & G. Chader, eds.) *Pergamon Press, Oxford.* **9**: 399-425.

Masu Y, Nakayama K, Tamaki H, Harada Y, Kuno M, Nakanishi S (1987) cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature.* **329**(6142): 836-838.

Mataga N, Nagai N, Hensch TK (2002) Permissive proteolytic activity for visual cortical plasticity. *Proc Natl Acad Sci U S A.* **99**(11): 7717-7721.

Matos MF, Mukherjee K, Chen X, Rizo J, Südhof TC (2003) Evidence for SNARE zippering during Ca^{2+}-triggered exocytosis in PC12 cells. *Neuropharmacology.* **45**: 777-786.

Matsui Y, Kikuchi A, Kondo J, Hishida T, Teranishi Y, Takai Y (1988) Nucleotide and deduced amino acid sequences of a GTP-binding protein family with molecular weights of 25,000 from bovine brain. *J Biol Chem.* **263**(23): 11071-11074.

Matsumoto-Miyai K, Ninomiya A, Yamasaki H, Tamura H, Nakamura Y, Shiosaka S (2003) NMDA-dependent proteolysis of presynaptic adhesion molecule L1 in the hippocampus by neuropsin. *J Neurosci.* **23**(21):7727-7736.

Matsuo I, Hirooka Y, Hironaga K, Eshima K, Shigematsu H, Shihara M, Sakai K, Takeshita A (2001) Glutamate release via NO production evoked by NMDA in the NTS enhances hypotension and bradycardia in vivo. *Am J Physiol Regul Integr Comp Physiol.* **280**(5): R1285-1291.

Matucz E, Moricz K, Gigler G, Simo A, Barkoczy J, Levay G, Harsing LG Jr, Szenasi G (2004) Reduction of cerebral infarct size by non-competitive AMPA antagonists in rats subjected to permanent and transient focal ischemia. *Brain Res.* **1019**(1-2): 210-216.

Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**(1): 374-378.

Mayer ML, Westbrook GL, Guthrie PB (1984) Voltage-dependent block by Mg2+ of NMDA responses in spinal cord neurones. *Nature.* **309**(5965): 261-263.

McGlade-McCulloh E, Yamamoto H, Tan SE, Brickey DA, Soderling TR (1993) Phosphorylation and regulation of glutamate receptors by calcium/calmodulin-dependent protein kinase II. *Nature.* **362**(6421): 640-642.

McKenna NJ, O'Malley BW (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell.* **108**(4): 465-474.

McKune K, Moore PA, Hull MW, Woychik NA (1995) Six human RNA polymerase subunits functionally substitute for their yeast counterparts. *Mol Cell Biol.* **15**(12): 6895-6900.

Melcher T, Maas S, Herb A, Sprengel R, Seeburg PH, Higuchi M (1996) A mammalian RNA editing enzyme. *Nature.* **379**(6564): 460-464.

Meldrum B, Garthwaite J (1990) Excitatory amino acid neurotoxicity and neurodegenerative disease. *Trends Pharmacol Sci.* **11**(9): 379-387.

Mellado M, Rodriguez-Frade JM, Aragay A, del Real G, Martin AM, Vila-Coro AJ, Serrano A, Mayor F Jr, Martinez-A C (1998) The chemokine monocyte chemotactic protein 1 triggers Janus kinase 2 activation and tyrosine phosphorylation of the CCR2B receptor. *J Immunol.* **161**(2): 805-813.

Mercurio F, Zhu H, Murray BW, Shevchenko A, Bennett BL, Li J, Young DB, Barbosa M, Mann M, Manning A, Rao A (1997) IKK-1 and IKK-2: cytokine-activated IkappaB kinases essential for NF-kappaB activation. *Science.* **278**(5339): 860-866.

Meyers EN, Lewandoski M, Martin GR (1998) An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. *Nat Genet.* **18**(2): 136-141.

Miyatake R, Furukawa A, Suwaki H (2002) Identification of a novel variant of the human NR2B gene promoter region and its possible association with schizophrenia. *Mol Psychiatry.* **7**(10): 1101-1106.

Mochida S (2000) Protein-protein interactions in neurotransmitter release. *Neurosci Res.* **36**(3): 175-182.

Molnar E, Baude A, Richmond SA, Patel PB, Somogyi P, McIlhinney RA (1993) Biochemical and immunocytochemical characterization of antipeptide antibodies to a cloned GluR1 glutamate receptor subunit: cellular and subcellular distribution in the rat forebrain. *Neuroscience.* **53**(2): 307-326.

Molnar E, Pickard L, Duckworth JK (2002) Developmental changes in ionotropic glutamate receptors: lessons from hippocampal synapses. *Neuroscientist.* **8**(2): 143-153.

Monaghan DT, Bridges RJ, Cotman CW (1989) The excitatory amino acid receptors: their classes, pharmacology, and distinct properties in the function of the central nervous system. *Annu Rev Pharmacol Toxicol.* **29**: 365-402.

Monyer H, Seeburg PH, Wisden W (1991) Glutamate-operated channels: developmentally early and mature forms arise by alternative splicing. *Neuron.* **6**(5): 799-810.

Morgan A, Burgoyne RD (1995) Is NSF a fusion protein? *Trends Cell Biol.* **5**(9): 335-339.

Moss SJ, Blackstone CD, Huganir RL (1993) Phosphorylation of recombinant non-NMDA glutamate receptors on serine and tyrosine residues. *Neurochem Res.* **18**(1): 105-110.

Moya KL, Tavitian B, Zahraoui A, Tavitian A. (1992) Localization of the ras-like rab3A protein in the adult rat brain. *Brain Res.* **590**(1-2): 118-127.

Myers SJ, Peters J, Huang Y, Comer MB, Barthel F, Dingledine R (1998) Transcriptional regulation of the GluR2 gene: neural-specific expression, multiple promoters, and regulatory elements. *J Neurosci.* **18**(17): 6723-6739.

Myers LC, Kornberg RD (2000) Mediator of transcriptional regulation. *Annu Rev Biochem.* **69**: 729-749.

Nagase T, Nakayama M, Nakajima D, Kikuno R, Ohara O (2001) Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.* **8**: 85-95.

Naka T, Narazaki M, Hirata M, Matsumoto T, Minamoto S, Aono A, Nishimoto N, Kajita T, Taga T, Yoshizaki K, Akira S, Kishimoto T (1997) Structure and function of a new STAT-induced STAT inhibitor. *Nature.* **387**(6636): 924-929.

Nakanishi N, Shneider NA, Axel R (1990) A family of glutamate receptor genes: evidence for the formation of heteromultimeric receptors with distinct channel properties. *Neuron.* **5**(5): 569-581.

Nakanishi S (1992) Molecular diversity of glutamate receptors and implications for brain function. *Science.* **258**(5082): 597-603.

Ng SK, Zhang Z, Tan SH, Lin K (2003) InterDom: a database of putative interacting protein domains for validating predicted protein interactions and complexes. *Nucleic Acids Res.* **31**(1): 251-254.

Nguyen VT, Benveniste EN. Involvement of STAT-1 and ets family members in interferon-gamma induction of CD40 transcription in microglia/macrophages. *J Biol Chem.* **275**(31): 23674-23684.

Nicholson SE, Willson TA, Farley A, Starr R, Zhang JG, Baca M, Alexander WS, Metcalf D, Hilton DJ, Nicola NA (1999) Mutational analyses of the SOCS proteins suggest a dual domain requirement but distinct mechanisms for inhibition of LIF and IL-6 signal transduction. *EMBO J.* **18**(2): 375-385.

Nicoll RA, Frerking M, Schmitz D (2000) AMPA receptors jump the synaptic cleft. *Nat Neurosci.* **3**(6):527-529.

Nicoll RA (2003) Expression mechanisms underlying long-term potentiation: a postsynaptic view. *Philos Trans R Soc Lond B Biol Sci.* **358**(1432): 721-726.

Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature.* **377**(6545): 119-128.

Nitta T (1998) Cytokine gene expression within the central nervous system. *Cell Mol Neurobiol.* **18**(6): 703-708.

Novick P, Zerial M (1997) The diversity of Rab proteins in vesicle transport. *Curr Opin Cell Biol.* **9**(4): 496-504.

Nowak L, Bregestovski P, Ascher P, Herbet A, Prochiantz A (1984) Magnesium gates glutamate-activated channels in mouse central neurones. *Nature.* **307**(5950): 462-465.

Oelgeschlager T, Chiang CM, Roeder RG (1996) Topology and reorganization of a human TFIID-promoter complex. *Nature.* **382**(6593): 735-738.

Ohara O, Nagase T, Ishikawa K, Nakajima D, Ohira M, Seki N, Nomura N (1997) Construction and characterization of human brain cDNA libraries suitable for analysis of cDNA clones encoding relatively large proteins. *DNA Res.* **4**(1): 53-59.

Ohnishi H, Ernst SA, Wys N, McNiven M, Williams JA (1996) Rab3D localizes to zymogen granules in rat pancreatic acini and other exocrine glands. *Am J Physiol.* **271**(3 Pt 1): G531-G538.

Oka T, Akisada M, Okabe A, Sakurai K, Shiosaka S, Kato K (2002) Extracellular serine protease neuropsin (KLK8) modulates neurite outgrowth and fasciculation of mouse hippocampal neurons in culture. *Neurosci Lett.* **321**(3): 141-144.

Orphanides G, Lagrange T, Reinberg D (1996) The general transcription factors of RNA polymerase II. *Genes Dev.* **10**(21): 2657-2683.

Ortiz J, Fitzgerald LW, Charlton M, Lane S, Trevisan L, Guitart X, Shoemaker W, Duman RS, Nestler EJ (1995) Biochemical actions of chronic ethanol exposure in the mesolimbic dopamine system. *Synapse.* **21**(4): 289-298.

Owens-Grillo JK, Hoffmann K, Hutchison KA, Yem AW, Deibel MR Jr, Handschumacher RE, Pratt WB (1995) The cyclosporin A-binding immunophilin CyP-40 and the FK506-binding immunophilin hsp56 bind to a common site on hsp90 and exist in independent cytosolic heterocomplexes with the untransformed glucocorticoid receptor. *J Biol Chem.* **270**(35): 20479-20484.

Owens-Grillo JK, Czar MJ, Hutchison KA, Hoffmann K, Perdew GH, Pratt WB (1996) A model of protein targeting mediated by immunophilins and other proteins that bind to hsp90 via tetratricopeptide repeat domains. *J Biol Chem.* **271**(23): 13468-13475.

Pardee TS, Bangur CS, Ponticelli AS (1998) The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. *J Biol Chem.* **273**(28): 17859-17864.

Park E, Velumian AA, Fehlings MG (2004) The role of excitotoxicity in secondary mechanisms of spinal cord injury: a review with an emphasis on the implications for white matter degeneration. *J Neurotrauma.* **21**(6):754-774.

Partin KM, Patneau DK, Winters CA, Mayer ML, Buonanno A (1993) Selective modulation of desensitization at AMPA versus kainate receptors by cyclothiazide and concanavalin A. *Neuron.* **11**(6): 1069-1082.

Passafaro M, Piech V, Sheng M (2001) Subunit-specific temporal and spatial patterns of AMPA receptor exocytosis in hippocampal neurons. *Nat Neurosci.* 4(9): 917-926.

Patikoglou GA, Kim JL, Sun L, Yang SH, Kodadek T, Burley SK (1999) TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* **13**(24): 3217-3230.

Patneau DK, Vyklicky L Jr, Mayer ML (1993) Hippocampal neurons exhibit cyclothiazide-sensitive rapidly desensitizing responses to kainate. *J Neurosci.* **13**(8): 3496-3509.

Pavlos NJ, Xu J, Papadimitriou JM, Zheng MH (2001) Molecular cloning of the mouse homologue of Rab3c. *J Mol Endocrinol.* **27**(1): 117-122.

Pedersen KM, Finsen B, Celis JE, Jensen NA (1999) muFKBP38: a novel murine immunophilin homolog differentially expressed in Schwannoma cells and central nervous system neurons in vivo. *Electrophoresis.* **20**(2): 249-255.

Pellegrini-Giampietro DE, Bennett MVL, Zukin RS (1991) Differential expression of three glutamate receptor genes in developing rat brain: an *in situ* hybridization study. *Proc. Natl. Acad. Sci., U.S.A.* **88**: 4157-4161.

Pellegrini-Giampietro DE, Bennett MVL, Zukin RS (1991) Differential expression of three glutamate receptor genes in developing rat brain: an *in situ* hybridization study. *Proc. Natl. Acad. Sci., U.S.A.* **88**: 4157-4161.

Pellegrini-Giampietro DE, Bennett MVL, Zukin RS (1992a) Are Ca^{2+}-permeable kainate/AMPA receptors more abundant in immature brain? *Neurosci. Lett.* **144**: 65-69.

Pellegrini-Giampietro DE, Zukin RS, Bennett MV, Cho S, Pulsinelli WA (1992b) Switch in glutamate receptor subunit gene expression in CA1 subfield of hippocampus following global ischemia in rats. *Proc. Natl. Acad. Sci. U.S.A.* **89**(21): 10499-10503.

Pellegrino M, Provero P, Silengo L, Di Cunto F (2004) CLOE: identification of putative functional relationships among genes by comparison of expression profiles between two species. *BMC Bioinformatics* **5**: 179.

Pennuto M, Dunlap D, Contestabile A, Benfenati F, Valtorta F (2002) Fluorescence resonance energy transfer detection of synaptophysin I and vesicle-associated membrane protein 2 interactions during exocytosis from single live synapses. *Mol Biol Cell.* **13**(8): 2706-2717.

Pennuto M, Bonanomi D, Benfenati F, Valtorta F (2003) Synaptophysin I controls the targeting of VAMP2/synaptobrevin II to synaptic vesicles. *Mol Biol Cell.* **14**(12): 4909-4919.

Perkinton MS, Sihra TS, Williams RJ (1999) Ca(2+)-permeable AMPA receptors induce phosphorylation of cAMP response element-binding protein through a phosphatidylinositol 3-kinase-dependent stimulation of the mitogen-activated protein kinase signaling cascade in neurons. *J Neurosci.* **19**(14): 5861-5874.

Pieri I, Klein M, Bayertz C, Gerspach J, van der Ploeg A, Pfizenmaier K, Eisel U (1999) Regulation of the murine NMDA-receptor-subunit NR2C promoter by Sp1 and fushi tarazu factor1 (FTZ-F1) homologues. *Eur J Neurosci.* **11**(6): 2083-2092.

Poirier MA, Xiao W, Macosko JC, Chan C, Shin YK, Bennett MK (1998) The synaptic SNARE complex is a parallel four-stranded helical bundle. *Nat Struct Biol.* **5**(9): 765-769.

Poisbeau P, Jo YH, Feltz P, Schlichter R (1996) Electrophysiological characterization of non-NMDA glutamate receptors on cultured intermediate lobe cells of the rat pituitary. *Neuroendocrinology.* **64**(2): 162-168.

Polizzotto MN, Bartlett PF, Turnley AM (2000) Expression of "suppressor of cytokine signalling" (SOCS) genes in the developing and adult mouse nervous system. *J Comp Neurol.* **423**(2): 348-358.

Pollard H, Heron A, Moreau J, Ben-Ari Y, Khrestchatisky M (1993) Alterations of the GluR-B AMPA receptor subunit flip/flop expression in kainate-induced epilepsy and ischemia. *Neuroscience.* **57**(3): 545-554.

Pratt WB, Czar MJ, Stancato LF, Owens JK (1993) The hsp56 immunophilin component of steroid receptor heterocomplexes: could this be the elusive nuclear localization signal-binding protein? *J Steroid Biochem Mol Biol.* **46**(3): 269-279.

Praz V, Périer R, Bonnard C, Bucher P (2002) The Eukaryotic Promoter Database, EPD: new entry types and links to gene expression data. *Nucleic Acids Res.* **30**(1): 322-324.

Price M.N, Rieffel E (2004) Finding coexpressed genes in counts-based data: an improved measure with validation experiments. *Bioinformatics.* **20**(6): 945-952.

Pruitt, K.D., Katz, K.S., Sicotte, H., Maglott, D.R. (2000) Introducing RefSeq and LocusLink: curated human genome resources at the NCBI. *Trends Genet.* **16**(1): 44-47.

Pruitt, K.D. and Maglott, D.R. (2001) RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.,* **29**(1): 137-140.

Quandt K, Frech K, Karas H, Wingender E, Werner T (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**(23): 4878-4884.

Qureshi SA, Jackson SP (1998) Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength. *Mol Cell.* **1**(3): 389-400.

Radanyi C, Chambraud B, Baulieu EE (1994) The ability of the immunophilin FKBP59-HBI to interact with the 90-kDa heat shock protein is encoded by its tetratricopeptide repeat domain. *Proc Natl Acad Sci U S A.* **91**(23): 11197-11201.

Rajan I, Cline HT (1998) Glutamate receptor activity is required for normal development of tectal cell dendrites in vivo. *J Neurosci.* **18**(19): 7836-7846.

Raman IM, Tong G, Jahr CE (1996) Beta-adrenergic regulation of synaptic NMDA receptors by cAMP-dependent protein kinase. *Neuron.* **16**(2): 415-421.

Ratajczak T, Carrello A (1996) Cyclophilin 40 (CyP-40), mapping of its hsp90 binding domain and evidence that FKBP52 competes with CyP-40 for hsp90 binding. *J Biol Chem.* **271**(6): 2961-2965.

Regazzi R, Kikuchi A, Takai Y, Wollheim CB (1992) The small GTP-binding proteins in the cytosol of insulin-secreting cells are complexed to GDP dissociation inhibitor proteins. *J Biol Chem.* **267**(25): 17512-17519.

Regnier CH, Song HY, Gao X, Goeddel DV, Cao Z, Rothe M (1997) Identification and characterization of an IkappaB kinase. *Cell.* **90**(2): 373-383.

Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA (2000) Genome-wide location and function of DNA binding proteins. *Science.* **290**(5500):2306-9.

Ren X, Carpenter A, Hogaboam C, Colletti L (2003) Mitogenic properties of endogenous and pharmacological doses of macrophage inflammatory protein-2 after 70% hepatectomy in the mouse. *Am J Pathol.* **163**(2): 563-570.

Renoir JM, Radanyi C, Faber LE, Baulieu EE (1990) The non-DNA-binding heterooligomeric form of mammalian steroid hormone receptors contains a hsp90-bound 59-kilodalton protein. *J Biol Chem.* **265**(18): 10740-10745

Rice LM, Brünger AT (1999) Crystal structure of the vesicular transport protein Sec17: implications for SNAP function in SNARE complex disassembly. Mol *Cell.* **4**(1): 85-95.

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16**(6): 276-277.

Robinson MS, Bonifacino JS (2001) Adaptor-related proteins. *Curr Opin Cell Biol.* **13**(4): 444-453.

Roche KW, O'Brien RJ, Mammen AL, Bernhardt J, Huganir RL (1996) Characterization of multiple phosphorylation sites on the AMPA receptor GluR1 subunit. *Neuron.* **16**(6): 1179-1188.

Rodriguez-Frade JM, Vila-Coro AJ, de Ana AM, Albar JP, Martinez-A C, Mellado M (1999a) The chemokine monocyte chemoattractant protein-1 induces functional responses through dimerization of its receptor CCR2. *Proc Natl Acad Sci U S A.* **96**(7): 3628-3633.

Rodriguez-Frade JM, Vila-Coro AJ, Martin A, Nieto M, Sanchez-Madrid F, Proudfoot AE, Wells TN, Martinez-A C, Mellado M (1999b) Similarities and differences in RANTES- and (AOP)-RANTES-triggered signals: implications for chemotaxis. *J Cell Biol.* **144**(4):755-765.

Rogers SW, Hughes TE, Hollmann M, Gasic GP, Deneris ES, Heinemann S (1991) The characterization and localization of the glutamate receptor subunit GluR1 in the rat brain. *J Neurosci.* **11**(9):2713-2724.

Rosenmund C, Carr DW, Bergeson SE, Nilaver G, Scott JD, Westbrook GL (1994) Anchoring of protein kinase A is required for modulation of AMPA/kainate receptors on hippocampal neurons. *Nature.* **368**(6474): 853-856.

Roy R, Schaeffer L, Humbert S, Vermeulen W, Weeda G, Egly JM (1994) The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH. *J Biol Chem.* **269**(13): 9826-9832.

Russell LC, Whitt SR, Chen MS, Chinkers M (1999) Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. *J Biol Chem.* **274**(29): 20060-20063.

Sakimura K, Bujo H, Kushiya E, Araki K, Yamazaki M, Yamazaki M, Meguro H, Warashina A, Numa S, Mishina M (1990) Functional expression from cloned cDNAs of glutamate receptor species responsive to kainate and quisqualate. *FEBS Lett.* **272**(1-2): 73-80.

Sakisaka T, Meerlo T, Matteson J, Plutner H, Balch WE (2002) Rab-alphaGDI activity is regulated by a Hsp90 chaperone complex. *EMBO J.* **21**(22): 6125-6135.

Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.* **32** (Database issue): D449-451.

Satake S, Saitow F, Yamada J, Konishi S (2000) Synaptic activation of AMPA receptors inhibits GABA release from cerebellar interneurons. *Nat Neurosci.* **3**(6): 551-558.

Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, Chambon P, Egly JM (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science.* **260**(5104): 58-63.

Schenk U, Verderio C, Benfenati F, Matteoli M (2003) Regulated delivery of AMPA receptor subunits to the presynaptic membrane. *EMBO J.* **22**(3): 558-568.

Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, Bartunik H, Hartl FU, Moarefi I (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell.* **101**(2): 199-210.

Schilling K, Gratzl M (1988) Quantification of p38/synaptophysin in highly purified adrenal medullary chromaffin vesicles. *FEBS Lett.* **233**(1): 22-24.

Schlüter OM, Khvotchev M, Jahn R, Südhof TC (2002) Localization versus function of Rab3 proteins. Evidence for a common regulatory role in controlling fusion. *J Biol Chem.* **277**(43): 40919-40929.

Schmid CD, Praz V, Delorenzi M, PérierR, Bucher P (2004) The Eukaryotic Promoter Database EPD: the impact of in silico primer extension. *Nucleic Acids Res.* **32**(Database issue): D82-85.

Schmidt KF, Kruse M, Hatt H (1994) Dopamine alters glutamate receptor desensitization in retinal horizontal cells of the perch (Perca fluviatilis). *Proc Natl Acad Sci U S A.* **91**(17): 8288-8291.

Schmidt C, Peng B, Li Z, Sclabas GM, Fujioka S, Niu J, Schmidt-Supprian M, Evans DB, Abbruzzese JL, Chiao PJ (2003) Mechanisms of proinflammatory cytokine-induced biphasic NF-kappaB activation. *Mol Cell.* **12**(5): 1287-1300.

Schneider A, Martin-Villalba A, Weih F, Vogel J, Wirth T, Schwaninger M (1999) NF-kappaB is activated and promotes cell death in focal cerebral ischemia. *Nat Med.* **5**(5): 554-559.

Schoenherr CJ, Anderson DJ (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science.* **267**(5202): 1360-1363.

Scholl T, Stevens MB, Mahanta S, Strominger JL (1996) A zinc finger protein that represses transcription of the human MHC class II gene, DPA. *J Immunol.* **156**(4): 1448-1457.

Schreiber SL (1991) Chemistry and biology of the immunophilins and their immunosuppressive ligands. *Science.* **251**(4991): 283-287.

Schuman EM, Madison DV (1994) Locally distributed synaptic potentiation in the hippocampus. *Science.* **263**(5146): 532-536.

Scott ML, Fujita T, Liou HC, Nolan GP, Baltimore D (1993) The p65 subunit of NF-kappa B regulates I kappa B by two distinct mechanisms. *Genes Dev.* **7**(7A): 1266-1276.

Seaman MN, Sowerby PJ, Robinson MS (1996) Cytosolic and membrane-associated proteins involved in the recruitment of AP-1 adaptors onto the trans-Golgi network. *J Biol Chem.* **271**(41): 25446-25451

Seeburg PH (1993) The TINS/TiPS Lecture. The molecular biology of mammalian glutamate receptor channels. *Trends Neurosci.* **16**(9): 359-365.

Seeburg PH (1996) The role of RNA editing in controlling glutamate receptor channel properties. *J Neurochem.* **66**(1): 1-5.

Serizawa H, Conaway RC, Conaway JW (1993) Multifunctional RNA polymerase II initiation factor delta from rat liver. Relationship between carboxyl-terminal domain kinase, ATPase, and DNA helicase activities. *J Biol Chem.* **268**(23): 17300-17308.

Shah NP, Witte ON, Denny CT (1991) Characterization of the BCR promoter in Philadelphia chromosome-positive and -negative cell lines. *Mol. Cell Biol.* **11**(4): 1854-1860.

Sharan R, Ben-Hur A, Loots GG, Ovcharenko I (2004) CREME: Cis-Regulatory Module Explorer for the human genome. *Nucleic Acids Res.* **32**(Web Server issue): W253-256.

Shelton MK, McCarthy KD (1999) Mature hippocampal astrocytes exhibit functional metabotropic and ionotropic glutamate receptors in situ. *Glia.* **26**(1): 1-11.

Shi S-H, Hayashi Y, Petralia RS, Zaman SH, Wenthold RJ, Svoboda K, Malinow R (1999) Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science.* **284**:1811-1816.

Shi S, Hayashi Y, Esteban JA, Malinow R (2001) Subunit-specific rules governing AMPA receptor trafficking to synapses in hippocampal pyramidal neurons. *Cell.* **105**(3): 331-343.

Shimojo M, Hersh LB (2004) Regulation of the cholinergic gene locus by the repressor element-1 silencing transcription factor/neuron restrictive silencer factor (REST/NRSF). *Life Sci.* **74**(18): 2213-2225.

Shiosaka S, Yoshida S (2000) Synaptic microenvironments--structural plasticity, adhesion molecules, proteases and their inhibitors. *Neurosci Res.* **37**(2):85-89

Shworak, N. W., Fritze, L. M. S., Liu, J., Butler, L. D., and Rosenberg, R. D. (1996) Cell-free synthesis of anticoagulant heparan sulfate reveals a limiting converting activity that modifies an excess precursor pool. *J. Biol. Chem.* **271**(43): 27063-27071.

Shworak NW, Liu J, Petros LM, Zhang L, Kobayashi M, Copeland NG, Jenkins NA, Rosenberg RD (1999) Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cdnas and identification of distinct genomic loci. *J Biol Chem.* **274**(8): 5170-5184.

Sidhu RS, Bhullar RP (2001) Rab3B in human platelet is membrane bound and interacts with Ca(2+)/calmodulin. *Biochem Biophys Res Commun.* **289**(5): 1039-1043.

Sigal NH, Dumont FJ (1992) Cyclosporin A, FK-506, and rapamycin: pharmacologic probes of lymphocyte signal transduction. *Annu Rev Immunol.* **10**: 519-560.

Silverstein AM, Galigniana MD, Kanelakis KC, Radanyi C, Renoir JM, Pratt WB (1999) Different regions of the immunophilin FKBP52 determine its association with the glucocorticoid receptor, hsp90, and cytoplasmic dynein. *J Biol Chem.* **274**(52): 36980-36986.

Sin WC, Haas K, Ruthazer ES, Cline HT (2002) Dendrite growth increased by visual activity requires NMDA receptor and Rho GTPases. *Nature.* **419**(6906): 475-480.

Skolnick P, Layer RT, Popik P, Nowak G, Paul IA, Trullas R (1996) Adaptation of N-methyl-D-aspartate (NMDA) receptors following antidepressant treatment: implications for the pharmacotherapy of depression. *Pharmacopsychiatry.* **29**(1): 23-26.

Smith AD, Sumazin P, Zhang MQ. (2005) Identifying tissue-selective transcription factor binding sites in vertebrate promoters. *Proc Natl Acad Sci U S A.* **102**(5):1560-5.

Smith TF, Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol.* **147**(1): 195-197.

Sollner T, Whiteheart SW, Brunner M, Erdjument-Bromage H, Geromanos S, Tempst P, Rothman JE (1993) SNAP receptors implicated in vesicle targeting and fusion. *Nature.* **362**(6418): 318-324.

Sommer B, Keinänen K, Verdoorn TA, Wisden W, Burnashev N, Herb A, Köhler M, Takagi T, Sakmann B, Seeburg PH (1990) Flip and flop: a cell-specific functional switch in glutamate-operated channels of the CNS. *Science.* **249**(4976): 1580-1585.

Sommer B, Köhler M, Sprengel R, Seeburg PH (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell.* **67**(1): 11-19.

Soriano SF, Serrano A, Hernanz-Falcon P, Martin de Ana A, Monterrubio M, Martinez C, Rodriguez-Frade JM, Mellado M (2003) Chemokines integrate JAK/STAT and G-protein pathways during chemotaxis and calcium flux responses. *Eur J Immunol.* **33**(5): 1328-1333.

Starr R, Willson TA, Viney EM, Murray LJ, Rayner JR, Jenkins BJ, Gonda TJ, Alexander WS, Metcalf D, Nicola NA, Hilton DJ (1997) A family of cytokine-inducible inhibitors of signalling. *Nature*. **387**(6636): 917-921.

Steegmaier M, Klumperman J, Foletti DL, Yoo JS, Scheller RH (1999) Vesicle-associated membrane protein 4 is implicated in trans-Golgi network vesicle trafficking. *Mol Biol Cell.* **10**(6): 1957-1972.

Stegenga SL, Kalb RG (2001) Developmental regulation of N-methyl-D-aspartate- and kainate-type glutamate receptor expression in the rat spinal cord. *Neuroscience.* **105**(2): 499-507.

Steiner JP, Dawson TM, Fotuhi M, Glatt CE, Snowman AM, Cohen N, Snyder SH (1992) High brain densities of the immunophilin FKBP colocalized with calcineurin. *Nature.* **358**(6387): 584-587.

Steiner JP, Dawson TM, Fotuhi M, Snyder SH (1996) Immunophilin regulation of neurotransmitter release. *Mol Med.* **2**(3): 325-333.

Stern-Bach Y, Bettler B, Hartley M, Sheppard PO, O'Hara PJ, Heinemann SF (1994) Agonist selectivity of glutamate receptors is specified by two domains structurally related to bacterial amino acid-binding proteins. *Neuron.* **13**(6): 1345-1357.

Stettler O, Nothias F, Tavitian B, Vernier P (1995) Double in situ hybridization reveals overlapping neuronal populations expressing the low molecular weight GTPases Rab3a and Rab3b in Rat brain. *Eur J Neurosci.* **7**(4): 702-713.

Steward O (1997) mRNA localization in neurons: a multipurpose mechanism? *Neuron.* **18**(1): 9-12.

Stryer L (1995) Biochemistry . *Fourth Edition. W.H. Freeman & Company, N.Y.*

Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discovery of conserved genetic modules. *Science.* **302**(5643): 249-255

Sugiyama H, Ito I, Hirono C (1987) A new type of glutamate receptor linked to inositol phospholipid metabolism. *Nature.* **325**(6104): 531-533.

Sun SC, Ganchi PA, Ballard DW, Greene WC (1993) NF-kappa B controls expression of inhibitor I kappa B alpha: evidence for an inducible autoregulatory pathway. *Science.* **259**(5103): 1912-1915.

Supino-Rosin L, Yoshimura A, Yarden Y, Elazar Z, Neumann D (2000) Intracellular retention and degradation of the epidermal growth factor receptor, two distinct processes mediated by benzoquinone ansamycins. *J Biol Chem.* **275**(29): 21850-21855.

Sutton RB, Fasshauer D, Jahn R, Brunger AT (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution. *Nature.* **395**(6700): 347-353.

Suyang H, Phillips R, Douglas I, Ghosh S (1996) Role of unphosphorylated, newly synthesized I kappa B beta in persistent activation of NF-kappa B. *Mol Cell Biol.* **16**(10): 5444-5449.

Suzuki Y, Yamashita R, Nakai K, Sugano S (2002) DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. *Nucleic Acids Res.* **30**(1): 328-331.

Suzuki Y, Yoshitomo-Nakagawa K, Maruyama K, Suyama A, Sugano S (1997) Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. *Gene.* **200**(1-2): 149-156.

Syrokou A, Tzanakakis G, Tsegenidis T, Hjerpe A, Karamanos NK (1999) Effects of glycosaminoglycans on proliferation of epithelial and fibroblast human malignant mesothelioma cells: a structure-function relationship. *Cell Prolif.* **32**(2-3): 85-99.

Tan SE, Wenthold RJ, Soderling TR (1994) Phosphorylation of AMPA-type glutamate receptors by calcium/calmodulin-dependent protein kinase II and protein kinase C in cultured hippocampal neurons. *J Neurosci.* **14**(3 Pt 1): 1123-1129.

Tan S, Hunziker Y, Sargent DF, Richmond TJ (1996) Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature.* **381**(6578): 127-151.

Tanaka Y, Adams DH, Hubscher S, Hirano H, Siebenlist U, Shaw S (1993) T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta. *Nature.* **361**(6407): 79-82.

Tatusova TA, Madden TL (1999) BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett.* **174**(2): 247-250.

Thomas KM, Taylor L, Navarro J (1991a) The interleukin-8 receptor is encoded by a neutrophil-specific cDNA clone, F3R. *J. Biol. Chem.* **266** (23): 14839-14841.

Thomas KR, Musci TS, Neumann PE, Capecchi MR (1991b) Swaying is a mutant allele of the proto-oncogene Wnt-1. *Cell.* **67**(5): 969-976.

Thompson JE, Phillips RJ, Erdjument-Bromage H, Tempst P, Ghosh S (1995) I kappa B-beta regulates the persistent response in a biphasic activation of NF-kappa B. *Cell.* **80**(4): 573-582.

Toni N, Buchs PA, Nikonenko I, Bron CR, Muller D (1999) LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature.* **402**(6760): 421-425.

Touchot N, Chardin P, Tavitian A (1987) Four additional members of the ras gene superfamily isolated by an oligonucleotide strategy: molecular cloning of YPT-related cDNAs from a rat brain library. *Proc Natl Acad Sci U S A.* **84**(23): 8210-8214.

Traub LM, Kornfeld S (1997) The trans-Golgi network: a late secretory sorting station. *Curr Opin Cell Biol.* **9**(4): 527-533.

Travagli J, Letourneur M, Bertoglio J, Pierre J (2004) STAT6 and Ets-1 form a stable complex that modulates Socs-1 expression by interleukin-4 in keratinocytes. *J Biol Chem.* **279**(34): 35183-35192.

Tsai FT, Sigler PB (2000) Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.* **19**(1): 25-36.

Tseng MT, Chang CC (1999) Ultrastructural localization of hippocampal TNF-alpha immunoreactive cells in rats following transient global ischemia. *Brain Res.* **833**(1): 121-124.

Turpin P, Hay RT, Dargemont C (1999) Characterization of IkappaBalpha nuclear import pathway. *J Biol Chem.* **274**(10):6804-6812.

Tuvim MJ, Adachi R, Chocano JF, Moore RH, Lampert RM, Zera E, Romero E, Knoll BJ, Dickey BF (1999) Rab3D, a small GTPase, is localized on mast cell secretory granules and translocates to the plasma membrane upon exocytosis. *Am J Respir Cell Mol Biol.* **20**(1): 79-89.

Valay JG, Simon M, Dubois MF, Bensaude O, Facca C, Faye G (1995) The KIN28 gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rpb1p CTD. *J Mol Biol.* **249**(3): 535-544.

Valentijn JA, Sengupta D, Gumkowski FD, Tang LH, Konieczko EM, Jamieson JD (1996a) Rab3D localizes to secretory granules in rat pancreatic acinar cells. *Eur J Cell Biol.* **70**(1): 33-41.

Valentijn JA, Gumkowski FD, Jamieson JD (1996b) The expression pattern of rab3D in the developing rat exocrine pancreas coincides with the acquisition of regulated exocytosis. *Eur J Cell Biol.* **71**(2): 129-136.

Velasco M, Diaz-Guerra MJ, Martin-Sanz P, Alvarez A, Bosca L (1997) Rapid Up-regulation of IkappaBbeta and abrogation of NF-kappaB activity in peritoneal macrophages stimulated with lipopolysaccharide. *J Biol Chem.* **272**(37): 23025-23030.

Verdoorn TA, Burnashev N, Monyer H, Seeburg PH, Sakmann B (1991) Structural determinants of ion flow through recombinant glutamate receptor channels. *Science.* **252**(5013): 1715-1718.

Verrijzer CP, Yokomori K, Chen JL, Tjian R (1994) Drosophila TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science.* **264**(5161): 933-941.

Verrijzer CP, Chen JL, Yokomori K, Tjian R (1995) Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell.* **81**(7): 1115-1125.

Vikman KS, Hill RH, Backstrom E, Robertson B, Kristensson K (2003) Interferon-gamma induces characteristics of central sensitization in spinal dorsal horn neurons in vitro. *Pain.* **106**(3): 241-251.

Vila-Coro AJ, Rodriguez-Frade JM, Martin De Ana A, Moreno-Ortiz MC, Martinez-A C, Mellado M (1999) The chemokine SDF-1alpha triggers CXCR4 receptor dimerization and activates the JAK/STAT pathway. *FASEB J.* **13**(13): 1699-1710.

Villalobos C, Núñez L, Garcia-Sancho J (1996) Functional glutamate receptors in a subpopulation of anterior pituitary cells. *FASEB J.* **10**(5): 654-660.

Vogt L, Schrimpf SP, Meskenaite V, Frischknecht R, Kinter J, Leone DP, Ziegler U, Sonderegger P (2001) Calsyntenin-1, a proteolytically processed postsynaptic membrane protein with a cytoplasmic calcium-binding domain. *Mol Cell Neurosci.* **17**(1): 151-166.

Wang LY, Salter MW, MacDonald JF (1991) Regulation of kainate receptors by cAMP-dependent protein kinase and phosphatases. *Science.* **253**(5024): 1132-1135.

Wang Y, Durkin JP (1995) alpha-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid, but not N-methyl-D-aspartate, activates mitogen-activated protein kinase through G-protein beta gamma subunits in rat cortical neurons. *J Biol Chem.* **270**(39): 22783-22787.

Wang Y, Small DL, Stanimirovic DB, Morley P, Durkin JP (1997) AMPA receptor-mediated regulation of a Gi-protein in cortical neurons. *Nature.* **389**(6650): 502-504.

Ward BK, Allan RK, Mok D, Temple SE, Taylor P, Dornan J, Mark PJ, Shaw DJ, Kumar P, Walkinshaw MD, Ratajczak T (2002) A structure-based mutational analysis of cyclophilin 40 identifies key residues in the core tetratricopeptide repeat domain that mediate binding to Hsp90. *J Biol Chem.* **277**(43): 40799-40809.

Washbourne P, Schiavo G, Montecucco C (1995) Vesicle-associated membrane protein-2 (synaptobrevin-2) forms a complex with synaptophysin. *Biochem J.* **305** (Pt 3): 721-724.

Watt TJ, Doyle DF (2005) ESPSearch: a program for finding exact sequences and patterns in DNA, RNA, or protein. *Biotechniques.* **38**(1): 109-115.

Weber E, Jilling T, Kirk KL (1996) Distinct functional properties of Rab3A and Rab3B in PC12 neuroendocrine cells. *J Biol Chem.* **271**(12): 6963-6971.

Weideman CA, Netter RC, Benjamin LR, McAllister JJ, Schmiedekamp LA, Coleman RA, Pugh BF (1997) Dynamic interplay of TFIIA, TBP and TATA DNA. *J Mol Biol.* **271**(1): 61-75.

Weil PA, Luse DS, Segall J, Roeder RG (1979) Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell.* **18**(2): 469-484.

Weinzierl ROJ (1999) Mechanisms of gene expression: Structure, function and evolution of the basal transcriptional machinery. *Imperial College Press.*

Weis L, Reinberg D (1997) Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. *Mol Cell Biol.* **17**(6): 2973-2984.

Wellenreuther R (2001) Generating and Sequencing Full-length cDNAs of Novel Human Genes within the German cDNA Consortium. *in* 11th Annual Workshop: Beyond the Identification of Transcribed Sequences: Functional and Expression Analysis, Washington D.C.

Wenthold RJ, Petralia RS, Blahos J II, Niedzielski AS (1996) Evidence for multiple AMPA receptor complexes in hippocampal CA1/CA2 neurons. *J Neurosci.* **16**(6): 1982-1989.

Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, Rapp BA (2002) Database resources of the National Center for Biotechnology Information: 2002 update. *Nucleic Acids Res.* **30**(1): 13-16.

Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L (2003) Database resources of the National Center for Biotechnology. *Nucleic Acids Res.* **31**(1): 28-33.

Wiedenmann B, Franke WW (1985) Identification and localization of synaptophysin, an integral membrane glycoprotein of Mr 38,000 characteristic of presynaptic vesicles. *Cell.* **41**(3): 1017-1028.

Wiemann S, Weil B, Wellenreuther R, Gassenhuber J, Glassl S, Ansorge W, Bocher M, Blocker H, Bauersachs S, Blum H, Lauber J, Dusterhoft A, Beyer A, Kohrer K, Strack N, Mewes HW, Ottenwalder B, Obermaier B, Tampe J, Heubner D, Wambutt R, Korn B, Klein M, Poustka A (2001) Toward a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs. *Genome Res.* **11**(3): 422-435.

Wo ZG, Oswald RE (1994) Transmembrane topology of two kainate receptor subunits revealed by N-glycosylation. *Proc Natl Acad Sci U S A.* **91**(15): 7154-7158.

Wo ZG, Oswald RE (1995) A topological analysis of goldfish kainate receptors predicts three transmembrane segments. *J Biol Chem.* **270**(5): 2000-2009.

Wolner BS, Gralla JD (2001) TATA-flanking sequences influence the rate and stability of TATA-binding protein and TFIIB binding. *J Biol Chem.* **276**(9): 6260-6266.

Wong VK, Yam JW, Hsiao WL (2003) Cloning and characterization of the promoter region of the mouse frizzled-related protein 4 gene. *Biol Chem.* **384**(8): 1147-1154.

Wu C, Ghosh S (2003) Differential phosphorylation of the signal-responsive domain of IκBα and IκBβ by IκB kinases. *J Biol Chem.* **278**(34): 31980-31987.

Wyllie DJ, Nicoll RA (1994) A role for protein kinases and phosphatases in the Ca(2+)-induced enhancement of hippocampal AMPA receptor-mediated synaptic responses. *Neuron.* **13**(3): 635-643.

Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D (2000) DIP: the database of interacting proteins. *Nucleic Acids Res.* **28**(1): 289-291.

Xu W, Mimnaugh EG, Kim JS, Trepel JB, Neckers LM (2002) Hsp90, not Grp94, regulates the intracellular trafficking and stability of nascent ErbB2. *Cell Stress Chaperones.* **7**(1): 91-96.

Yakel JL, Vissavajjhala P, Derkach VA, Brickey DA, Soderling TR (1995) Identification of a Ca^{2+}/calmodulin-dependent protein kinase II regulatory phosphorylation site in non-N-methyl-D-aspartate glutamate receptors. *Proc Natl Acad Sci U S A.* **92**(5): 1376-1380.

Yamada KA, Tang CM (1993) Benzothiadiazides inhibit rapid glutamate receptor desensitization and enhance glutamatergic synaptic currents. *J Neurosci.* **13**(9): 3904-3915.

Yamaguchi Y (2001) Heparan sulfate proteoglycans in the nervous system: their diverse roles in neurogenesis, axon guidance, and synaptogenesis. *Semin Cell Dev Biol.* **12**(2): 99-106.

Yamashita R, Suzuki Y, Sugano S, Nakai K (2001) DBTSS : DataBase of Human Transcriptional Start Sites and Full-Length cDNA. *Genome Informatics.* **12**: 488-489.

Yasukawa H, Misawa H, Sakamoto H, Masuhara M, Sasaki A, Wakioka T, Ohtsuka S, Imaizumi T, Matsuda T, Ihle JN, Yoshimura A (1999) The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. *EMBO J.* **18**(5): 1309-1320.

Yasukawa H, Yajima T, Duplain H, Iwatate M, Kido M, Hoshijima M, Weitzman MD, Nakamura T, Woodard S, Xiong D, Yoshimura A, Chien KR, Knowlton KU (2003) The suppressor of cytokine signaling-1 (SOCS1) is a novel therapeutic target for enterovirus-induced cardiac injury. *J Clin Invest.* **111**(4): 469-478.

Yoshida Y, Kurosawa N, Kanematsu T, Kojima N, Tsuji S (1996) Genomic structure and promoter activity of the mouse polysialic acid synthase gene (mST8Sia II). Brain-specific expression from a TATA-less GC-rich sequence. *J Biol Chem.* **271**(47): 30167-30173.

Young JC, Obermann WM, Hartl FU (1998) Specific binding of tetratricopeptide repeat proteins to the C-terminal 12-kDa domain of hsp90. *J Biol Chem.* **273**(29): 18007-18010.

Yu RC, Hanson PI, Jahn R, Brünger AT (1998) Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat Struct Biol.* **5**(9): 803-811.

Yu RC, Jahn R, Brünger AT (1999) NSF N-terminal domain crystal structure: models of NSF function. *Mol Cell.* **4**(1): 97-107.

Yu ZF, Nikolova-Karakashian M, Zhou D, Cheng G, Schuchman EH, Mattson MP (2000) Pivotal role for acidic sphingomyelinase in cerebral ischemia-induced ceramide and cytokine production, and neuronal apoptosis. *J Mol Neurosci.* **15**(2): 85-97.

Yu Z, Cheng G, Wen X, Wu GD, Lee WT, Pleasure D (2002) Tumor necrosis factor alpha increases neuronal vulnerability to excitotoxic necrosis by inducing expression of the AMPA-glutamate receptor subunit GluR1 via an acid sphingomyelinase- and NF-kappaB-dependent mechanism. *Neurobiol Dis.* **11**(1):199-213.

Yu X, Zhu X, Pi W, Ling J, Ko L, Takeda Y, Tuan D (2005) The long terminal repeat (LTR) of ERV-9 human endogenous retrovirus binds to NF-Y in the assembly of an active LTR enhancer complex NF-Y/MZF1/GATA-2. *J Biol Chem.* **280**(42): 35184-35194.

Yuste R, Bonhoeffer T (2001) Morphological changes in dendritic spines associated with long-term synaptic plasticity. *Annu Rev Neurosci.* **24**: 1071-1089.
Zamanillo D, Sprengel R, Hvalby O, Jensen V, Burnashev N, Rozov A, Kaiser KM, Koster HJ, Borchardt T, Worley P, Lubke J, Frotscher M, Kelly PH, Sommer B, Andersen P, Seeburg PH, Sakmann B (1999) Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science.* **284**(5421): 1805-1811.

Zabel U, Henkel T, Silva MS, Baeuerle PA (1993) Nuclear uptake control of NF-kappa B by MAD-3, an I kappa B protein present in the nucleus. *EMBO J.* **12**(1): 201-211.

Zahraoui A, Touchot N, Chardin P, Tavitian A (1989) The human Rab genes encode a family of GTP-binding proteins related to yeast YPT1 and SEC4 products involved in secretion. *J Biol Chem.* **264**(21): 12394-12401.

Zamanillo D, Sprengel R, Hvalby O, Jensen V, Burnashev N, Rozov A, Kaiser KM, Koster HJ, Borchardt T, Worley P, Lubke J, Frotscher M, Kelly PH, Sommer B, Andersen P, Seeburg PH, Sakmann B (1999) Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science.* **284**(5421): 1805-1811.

Zhang T, Zhang M (2001) Promoter Extraction from GenBank (PEG): automatic extraction of eukaryotic promoter sequences in large sets of genes. Bioinformatics. **17**(12): 1232-1233.

Zhang H, Ramanathan Y, Soteropoulos P, Recce ML, Tolias PP (2002) EZ-Retrieve: a web-server for batch retrieval of coordinate-specified human DNA sequences and underscoring putative transcription factor-binding sites. *Nucleic Acids Res.* **30**(21):e121.

Zhao F, Xuan Z, Liu L, Zhang MQ (2005) TRED: a Transcriptional Regulatory Element Database and a platform for in silico gene regulation studies. *Nucleic Acids Res.* **33**(Database Issue): D103-107.

Zinsmaier KE, Bronk P (2001) Molecular chaperones and the regulation of neurotransmitter exocytosis. *Biochem Pharmacol.* **62**(1): 1-11.

For Table 4.2:

Advani RJ, Bae HR, Bock JB, Chao DS, Doung YC, Prekeris R, Yoo JS, Scheller RH (1998) Seven novel mammalian SNARE proteins localize to distinct membrane compartments. *J Biol Chem.* **273**(17): 10317-10324.

Agrawal DK, Hopfenspirger MT, Chavez J, Talmadge JE (2001) Flt3 ligand: a novel cytokine prevents allergic asthma in a mouse model. *Int Immunopharmacol.* **1**(12): 2081-2089

Amaravadi R, Glerum DM, Tzagoloff A (1997) Isolation of a cDNA encoding the human homolog of COX17, a yeast gene essential for mitochondrial copper recruitment. *Hum Genet.* **99**(3): 329-333.

Anisowicz A, Sotiropoulou G, Stenman G, Mok SC, Sager R (1996) A novel protease homolog differentially expressed in breast and ovarian cancer. *Mol Med.* **2**(5): 624-636.

Baier M, Bannert N, Werner A, Lang K, Kurth R (1997) Molecular cloning, sequence, expression, and processing of the interleukin 16 precursor. *Proc Natl Acad Sci U S A*. **94**(10): 5273-5277.

Bi W, Yan J, Stankiewicz P, Park SS, Walz K, Boerkoel CF, Potocki L, Shaffer LG, Devriendt K, Nowaczyk MJ, Inoue K, Lupski JR (2002) Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. *Genome Res.* **12**(5): 713-728.

Chandrasekaran EV, Lakhaman SS, Chawda R, Piskorz CF, Neelamegham S, Matta KL (2004) Identification of physiologically relevant substrates for cloned Gal: 3-O-sulfotransferases (Gal3STs): distinct high affinity of Gal3ST-2 and LS180 sulfotransferase for the globo H backbone, Gal3ST-3 for N-glycan multiterminal Galbeta1, 4GlcNAcbeta units and 6-sulfoGalbeta1, 4GlcNAcbeta, and Gal3ST-4 for the mucin core-2 trisaccharide. *J Biol Chem.* **279**(11): 10032-10041.

Clark ME, Kelner GS, Turbeville LA, Boyer A, Arden KC, Maki RA (2000) ADAMTS9, a novel member of the ADAM-TS/ metallospondin gene family. *Genomics.* **67**(3): 343-350.

Clark HF, Gurney AL, Abaya E, Baker K, Baldwin D, Brush J, Chen J, Chow B, Chui C, Crowley C, Currell B, Deuel B, Dowd P, Eaton D, Foster J, Grimaldi C, Gu Q, Hass PE, Heldens S, Huang A, Kim HS, Klimowski L, Jin Y, Johnson S, Lee J, Lewis L, Liao D, Mark M, Robbie E, Sanchez C, Schoenfeld J, Seshagiri S, Simmons L, Singh J, Smith V, Stinson J, Vagts A, Vandlen R, Watanabe C, Wieand D, Woods K, Xie MH, Yansura D, Yi S, Yu G, Yuan J, Zhang M, Zhang Z, Goddard A, Wood WI, Godowski P, Gray A (2003) The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. *Genome Res.* **13**(10): 2265-2270.

Cruikshank WW, Center DM, Nisar N, Wu M, Natke B, Theodore AC, Kornfeld H (1994) Molecular and functional analysis of a lymphocyte chemoattractant factor: association of biologic function with CD4 expression. *Proc Natl Acad Sci U S A.* **91**(11): 5109-5113.

Dear TN, Moller A, Boehm T (1999) CAPN11: A calpain with high mRNA levels in testis and located on chromosome 6. *Genomics.* **59**(2): 243-247.

Eisenhaber B, Maurer-Stroh S, Novatchkova M, Schneider G, Eisenhaber F (2003) Enzymes and auxiliary factors for GPI lipid anchor biosynthesis and post-translational transfer to proteins. *Bioessays.* **25**(4): 367-385.

Espinosa L, Martin M, Nicolas A, Fabre M, Navarro E (1997) Primary sequence of the human, lysine-rich, ribosomal protein RPL38 and detection of an unusual RPL38 processed pseudogene in the promoter region of the type-1 angiotensin II receptor gene. *Biochim Biophys Acta.* **1354**(1): 58-64.

Ferri N, Carragher NO, Raines EW (2004) Role of discoidin domain receptors 1 and 2 in human smooth muscle cell-mediated collagen remodeling: potential implications in atherosclerosis and lymphangioleiomyomatosis. *Am J Pathol.* **164**(5): 1575-1585.

Fimia GM, De Cesare D, Sassone-Corsi P (1999) CBP-independent activation of CREM and CREB by the LIM-only protein ACT. *Nature.* **398**(6723): 165-169.

Fimia GM, Morlon A, Macho B, De Cesare D, Sassone-Corsi P (2001) Transcriptional cascades during spermatogenesis: pivotal role of CREM and ACT. *Mol Cell Endocrinol.* **179**(1-2): 17-23.

Garibay-Tupas JL, Csiszar K, Fox M, Povey S, Bryant-Greenwood GD (1999) Analysis of the 5'-upstream regions of the human relaxin H1 and H2 genes and their chromosomal localization on chromosome 9p24.1 by radiation hybrid and breakpoint mapping. *J Mol Endocrinol.* **23**(3): 355-365.

Hannum C, Culpepper J, Campbell D, McClanahan T, Zurawski S, Bazan JF, Kastelein R, Hudak S, Wagner J, Mattson J (1994) Ligand for FLT3/FLK2 receptor tyrosine kinase

regulates growth of haematopoietic stem cells and is encoded by variant RNAs. *Nature.* **368**(6472): 643-648.

Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park HS, Toyoda A, Ishii K, Totoki Y, Choi DK, Groner Y, Soeda E, Ohki M, Takagi T, Sakaki Y, Taudien S, Blechschmidt K, Polley A, Menzel U, Delabar J, Kumpf K, Lehmann R, Patterson D, Reichwald K, Rump A, Schillhabel M, Schudy A, Zimmermann W, Rosenthal A, Kudoh J, Schibuya K, Kawasaki K, Asakawa S, Shintani A, Sasaki T, Nagamine K, Mitsuyama S, Antonarakis SE, Minoshima S, Shimizu N, Nordsiek G, Hornischer K, Brant P, Scharfe M, Schon O, Desario A, Reichelt J, Kauer G, Blocker H, Ramser J, Beck A, Klages S, Hennig S, Riesselmann L, Dagand E, Haaf T, Wehrmeyer S, Borzym K, Gardiner K, Nizetic D, Francis F, Lehrach H, Reinhardt R, Yaspo ML; Chromosome 21 mapping and sequencing consortium (2000) The DNA sequence of human chromosome 21. *Nature.* **405**(6784): 311-319.

Hintsch G, Zurlinden A, Meskenaite V, Steuble M, Fink-Widmer K, Kinter J, Sonderegger P (2002) The calsyntenins--a family of postsynaptic membrane proteins with distinct neuronal expression patterns. *Mol Cell Neurosci.* **21**(3): 393-409.

Horvath R, Lochmuller H, Stucka R, Yao J, Shoubridge EA, Kim SH, Gerbitz KD, Jaksch M (2000) Characterization of human SCO1 and COX17 genes in mitochondrial cytochrome-c-oxidase deficiency. *Biochem Biophys Res Commun.* **276**(2): 530-533.

Hudson P, Haley J, John M, Cronk M, Crawford R, Haralambidis J, Tregear G, Shine J, Niall H (1983) Structure of a genomic clone encoding biologically active human relaxin. *Nature.* **301**(5901): 628-631.

Inoue N, Watanabe R, Takeda J, Kinoshita T (1996) PIG-C, one of the three human genes involved in the first step of glycosylphosphatidylinositol biosynthesis is a homologue of Saccharomyces cerevisiae GPI2. *Biochem Biophys Res Commun.* **226**(1): 193-199.

Karn T, Holtrich U, Brauninger A, Bohme B, Wolf G, Rubsamen-Waigmann H, Strebhardt K (1993) Structure, expression and chromosomal mapping of TKT from man and mouse: a new subclass of receptor tyrosine kinases with a factor VIII-like domain. *Oncogene.* **8**(12): 3433-3440.

Katoh M, Katoh M (2004) Identification and characterization of human FHOD3 gene in silico. *Int J Mol Med.* **13**(4): 615-620.

Kee HJ, Koh JT, Kim MY, Ahn KY, Kim JK, Bae CS, Park SS, Kim KK (2002) Expression of brain-specific angiogenesis inhibitor 2 (BAI2) in normal and ischemic brain: involvement of BAI2 in the ischemia-induced brain angiogenesis. *J Cereb Blood Flow Metab.* **22**(9): 1054-1067.

Ki CS, Na DL, Kim DK, Kim HJ, Kim JW (2002) Genetic association of an apolipoprotein C-I (APOC1) gene polymorphism with late-onset Alzheimer's disease. *Neurosci Lett.* **319**(2): 75-78.

Koc EC, Burkhart W, Blackburn K, Moyer MB, Schlatzer DM, Moseley A, Spremulli LL (2001) The large subunit of the mammalian mitochondrial ribosome. Analysis of the complement of ribosomal proteins present. *J Biol Chem.* **276**(47): 43958-43969.

Krebs DL, Hilton DJ (2000) SOCS: physiological suppressors of cytokine signaling. *J Cell Sci.* **113** (Pt 16): 2813-2819.

Krupnik VE, Sharp JD, Jiang C, Robison K, Chickering TW, Amaravadi L, Brown DE, Guyot D, Mays G, Leiby K, Chang B, Duong T, Goodearl AD, Gearing DP, Sokol SY, McCarthy SA (1999) Functional and structural diversity of the human Dickkopf gene family. *Gene.* **238**(2): 301-313.

Kurschner C, Yuzaki M (1999) Neuronal interleukin-16 (NIL-16): a dual function PDZ domain protein. *J Neurosci.* **19**(18): 7770-7780.

Lam E, Martin M, Wiederrecht G (1995) Isolation of a cDNA encoding a novel human FK506-binding protein homolog containing leucine zipper and tetratricopeptide repeat motifs. *Gene.* **160**(2): 297-302.

Law GL, Itoh H, Law DJ, Mize GJ, Merchant JL, Morris DR (1998) Transcription factor ZBP-89 regulates the activity of the ornithine decarboxylase promoter. *J Biol Chem.* **273**(32): 19955-19964.

Li Q, Verma IM (2002) NF-kappaB regulation in the immune system. *Nat Rev Immunol.* **2**(10): 725-734.

Li Y, Uruno T, Haudenschild C, Dudek SM, Garcia JG, Zhan X (2004) Interaction of cortactin and Arp2/3 complex is required for sphingosine-1-phosphate-induced endothelial cell remodeling. *Exp Cell Res.* **298**(1): 107-121.

Lin M, Napoli JL (2000) cDNA cloning and expression of a human aldehyde dehydrogenase (ALDH) active with 9-cis-retinal and identification of a rat ortholog, ALDH12. *J Biol Chem.* **275**(51): 40106-40112.

Liu J, Shworak NW, Sinay P, Schwartz JJ, Zhang L, Fritze LM, Rosenberg RD (1999a) Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities. *J Biol Chem.* **274**(8): 5185-5192.

Liu J, Shriver Z, Blaiklock P, Yoshida K, Sasisekharan R, Rosenberg RD (1999b) Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3A sulfates N-unsubstituted glucosamine residues. *J Biol Chem.* **274**(53): 38155-38162.

Lyman SD, James L, Johnson L, Brasel K, de Vries P, Escobar SS, Downey H, Splett RR, Beckmann MP, McKenna HJ (1994) Cloning of the human homologue of the murine flt3 ligand: a growth factor for early hematopoietic progenitor cells. *Blood.* **83**(10): 2795-2801.

Lynch DK, Winata SC, Lyons RJ, Hughes WE, Lehrbach GM, Wasinger V, Corthals G, Cordwell S, Daly RJ (2003) A Cortactin-CD2-associated protein (CD2AP) complex provides a novel link between epidermal growth factor receptor endocytosis and the actin cytoskeleton. *J Biol Chem.* **278**(24): 21805-21813.

Martin LH, Calabi F, Lefebvre FA, Bilsland CA, Milstein C (1987) Structure and expression of the human thymocyte antigens CD1a, CD1b, and CD1c. *Proc Natl Acad Sci U S A.* **84**(24): 9189-9193.

Mehta R, Gantz DL, Gursky O (2003) Effects of mutations in apolipoprotein C-1 on the reconstitution and kinetic stability of discoidal lipoproteins. *Biochemistry.* **42**(16): 4751-4758.

Nagase T, Ishikawa K, Suyama M, Kikuno R, Miyajima N, Tanaka A, Kotani H, Nomura N, Ohara O (1998) Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res.* **5**(5): 277-286.

Nagase T, Kikuno R, Hattori A, Kondo Y, Okumura K, Ohara O (2000) Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res.* **7**(6): 347-355.

Nagpal S, Patel S, Jacobe H, DiSepio D, Ghosn C, Malhotra M, Teng M, Duvic M, Chandraratna RA (1997) Tazarotene-induced gene 2 (TIG2), a novel retinoid-responsive gene in skin. *J Invest Dermatol.* **109**(1): 91-95.

Pan W, Zhang Q, Xi QS, Gan RB, Li TP (2001) FUP1, a gene associated with hepatocellular carcinoma, stimulates NIH3T3 cell proliferation and tumor formation in nude mice. *Biochem Biophys Res Commun.* **286**(5): 1033-1038.

Pedersen KM, Finsen B, Celis JE, Jensen NA (1999) muFKBP38: a novel murine immunophilin homolog differentially expressed in Schwannoma cells and central nervous system neurons in vivo. *Electrophoresis.* **20**(2): 249-255.

Punter FA, Adams DL, Glerum DM (2000) Characterization and localization of human COX17, a gene involved in mitochondrial copper transport. *Hum Genet.* **107**(1): 69-74.

Reymond A, Friedli M, Henrichsen CN, Chapot F, Deutsch S, Ucla C, Rossier C, Lyle R, Guipponi M, Antonarakis SE (2001) From PREDs and open reading frames to cDNA isolation: revisiting the human chromosome 21 transcription map. *Genomics.* **78**(1-2): 46-54.

Sampson EM, Haque ZK, Ku MC, Tevosian SG, Albanese C, Pestell RG, Paulson KE, Yee AS (2001) Negative regulation of the Wnt-beta-catenin pathway by the transcriptional repressor HBP1. *EMBO J.* **20**(16): 4500-4511.

Scarisbrick IA, Blaber SI, Lucchinetti CF, Genain CP, Blaber M, Rodriguez M (2002) Activity of a newly identified serine protease in CNS demyelination. *Brain.* **125**(Pt 6): 1283-1296.

Schlüter OM, Khvotchev M, Jahn R, Südhof TC (2002) Localization versus function of Rab3 proteins. Evidence for a common regulatory role in controlling fusion. *J Biol Chem.* **277**(43): 40919-40929.

Shiratsuchi T, Nishimori H, Ichise H, Nakamura Y, Tokino T (1997) Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI1). *Cytogenet Cell Genet.* **79**(1-2): 103-108.

Starr R, Willson TA, Viney EM, Murray LJ, Rayner JR, Jenkins BJ, Gonda TJ, Alexander WS, Metcalf D, Nicola NA, Hilton DJ (1997) A family of cytokine-inducible inhibitors of signalling. *Nature*. **387**(6636): 917-921.

Steegmaier M, Klumperman J, Foletti DL, Yoo JS, Scheller RH (1999) Vesicle-associated membrane protein 4 is implicated in trans-Golgi network vesicle trafficking. *Mol Biol Cell.* **10**(6): 1957-1972.

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ, Marra MA; Mammalian Gene Collection Program Team (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A.* **99**(26): 16899-16903.

Swanson KA, Knoepfler PS, Huang K, Kang RS, Cowley SM, Laherty CD, Eisenman RN, Radhakrishnan I (2004) HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. *Nat Struct Mol Biol.* **11**(8): 738-746.

Takash W, Canizares J, Bonneaud N, Poulat F, Mattei MG, Jay P, Berta P (2001) SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. *Nucleic Acids Res.* **29**(21): 4274-4283.

Tevosian SG, Shih HH, Mendelson KG, Sheppard KA, Paulson KE, Yee AS (1997) HBP1: a HMG box transcriptional repressor that is targeted by the retinoblastoma family. *Genes Dev.* **11**(3): 383-396.

Thompson JE, Phillips RJ, Erdjument-Bromage H, Tempst P, Ghosh S (1995) I kappa B-beta regulates the persistent response in a biphasic activation of NF-kappa B. *Cell.* **80**(4): 573-582.

Vogel W, Gish GD, Alves F, Pawson T (1997) The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol Cell.* **1**(1): 13-23.

Vogt L, Schrimpf SP, Meskenaite V, Frischknecht R, Kinter J, Leone DP, Ziegler U, Sonderegger P (2001) Calsyntenin-1, a proteolytically processed postsynaptic membrane protein with a cytoplasmic calcium-binding domain. *Mol Cell Neurosci.* **17**(1): 151-166.

Wang Y, Kobori JA, Hood L (1993) The ht beta gene encodes a novel CACCC box-binding protein that regulates T-cell receptor gene expression. *Mol Cell Biol.* **13**(9): 5691-5701.

Wang L, Darling J, Zhang JS, Liu W, Qian J, Bostwick D, Hartmann L, Jenkins R, Bardenhauer W, Schutte J, Opalka B, Smith DI (2000) Loss of expression of the DRR 1 gene at chromosomal segment 3p21.1 in renal cell carcinoma. *Genes Chromosomes Cancer.* **27**(1): 1-10.

Watanabe R, Inoue N, Westfall B, Taron CH, Orlean P, Takeda J, Kinoshita T (1998) The first step of glycosylphosphatidylinositol biosynthesis is mediated by a complex of PIG-A, PIG-H, PIG-C and GPI1. *EMBO J.* **17**(4): 877-885.

Wu Y, Diab I, Zhang X, Izmailova ES, Zehner ZE (2004) Stat3 enhances vimentin gene expression by binding to the antisilencer element and interacting with the repressor protein, ZBP-89. *Oncogene.* **23**(1): 168-178.

Yamato T, Orikasa K, Fukushige S, Orikasa S, Horii A (1999) Isolation and characterization of the novel gene, TU3A, in a commonly deleted region on 3p14.3-->p14.2 in renal cell carcinoma. *Cytogenet Cell Genet.* **87**(3-4): 291-295.

Yang QS, Xia F, Gu SH, Yuan HL, Chen JZ, Yang QS, Ying K, Xie Y, Mao YM (2002) Cloning and expression pattern of a spermatogenesis-related gene, BEX1, mapped to chromosome Xq22. *Biochem Genet.* **40**(1-2): 1-12.

Yarar D, D'Alessio JA, Jeng RL, Welch MD (2002) Motility determinants in WASP family proteins. *Mol Biol Cell.* **13**(11): 4045-4059.

Yoshihama M, Uechi T, Asakawa S, Kawasaki K, Kato S, Higa S, Maeda N, Minoshima S, Tanaka T, Shimizu N, Kenmochi N (2002) The human ribosomal protein genes: sequencing and comparative analysis of 73 genes. *Genome Res.* **12**(3): 379-390.

Yousef GM, Luo LY, Scherer SW, Sotiropoulou G, Diamandis EP (1999) Molecular characterization of zyme/protease M/neurosin (PRSS9), a hormonally regulated kallikrein-like serine protease. *Genomics.* **62**(2): 251-259.

Zahraoui A, Touchot N, Chardin P, Tavitian A (1989) The human Rab genes encode a family of GTP-binding proteins related to yeast YPT1 and SEC4 products involved in secretion. *J Biol Chem.* **264**(21): 12394-12401.

Zhang QH, Ye M, Wu XY, Ren SX, Zhao M, Zhao CJ, Fu G, Shen Y, Fan HY, Lu G, Zhong M, Xu XR, Han ZG, Zhang JW, Tao J, Huang QH, Zhou J, Hu GX, Gu J, Chen SJ, Chen Z (2000) Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells. *Genome Res.* **10**(10): 1546-1560.

Zhang Z, Gerstein M (2003) Identification and characterization of over 100 mitochondrial ribosomal protein pseudogenes in the human genome. *Genomics.* **81**(5): 468-480.

Zheng Z, Venkatapathy S, Rao G, Harrington CA (2002) Expression profiling of B cell chronic lymphocytic leukemia suggests deficient CD1-mediated immunity, polarized cytokine response, altered adhesion and increased intracellular protein transport and processing of leukemic cells. *Leukemia.* **16**(12): 2429-2437.

ABBREVIATIONS

Å	angstrom
AMPA	α-amino-3-hydroxy-5-methyl-isoxazole-4-propionate
ASMase	acid sphingomyelinase
AVA	associated valid accession
BDNF	brain-derived neurotrophic factor
BIND	Biomolecular Interaction Database
bp	base pair
BRE	TFIIB recognition element
CaM	calmodulin
CAMKII	calmodulin-dependent protein kinase II
CDS	protein coding sequence
CNQX	6-cyano-7-nitroquinoxaline-2,3-dione
CGN	cerebellar granule neuron
CNS	central nervous system
CsA	cyclosporin A
CyP40	cyclophilin 40
DIP	Database of Interacting Proteins
DPE	downstream promoter element
EMSA	electrophoretic mobility shift assay
EPD	Eukaryotic Promoter Database
EPSC	excitatory postsynaptic current
ER	endoplasmic reticulum
ERK	extracellular signal-regulated kinase
EV	Evidence Viewer (in NCBI's LocusLink)
FIE2	5'-end Information Extraction (version 2.0)
FKBP	FK506 binding protein
FN	false negative
FP	false positive
GDNF	glial-cell line derived neurotrophic factor
Glu	glutamate
GluR	glutamate receptors
GPCR	G-protein coupled receptor
GRIA	AMPA ionotropic glutamate receptor
GRIN	NMDA ionotropic glutamate receptor
GROß	gowth-related gene product ß
GTF	general transcription factor
HEK	human embryonic kidney
HS	heparan sulfate
hsp90	heat-shock protein 90
HUGO	Human Genome Organization
IFN-γ	interferon-gamma
IκB	inhibitor of NF-κB
IKK	IκB kinase
IL	interleukin

Inr	intiator element
JAK	Janus kinase
KA	kainate
KEGG	Kyoto Encyclopedia of Genes and Genomes
kDa	kiloDalton
LIF	leukemia-inhibitory factor
LPS	lipopolysaccharide
LTD	long-term depression
LTP	long-term potentiation
NES	nuclear-export signal
NF-κB	nuclear factor-kappa B
NLS	nuclear-localisation signal
NMDA	N-methyl-D-aspartate
NO	nitric oxide
nNOS	neuronal nitric oxide synthase
NSF	N-ethylmaleimide-sensitive fusion protein
nt.	nucleotide
MAPK	mitogen-activated protein kinase
MIP-2	macrophage inflamatory protein-2
NRF-1	nuclear respiratory factor-1
NRSE	neuron-restrictive silencer element
NRSF	neuron-restrictive silencer factor
PEG	Promoter Extraction from GenBank
PI3K	phosphatidylinositol 3-kinase / phosphoinositide 3-kinase
PIC	preinitiation complex
PKA	cAMP-dependent protein kinase
PKC	protein kinase C
PTX	pertussis toxin
QA	quisqualate
REST	RE1-silencing transcription factor
RNA Pol II	RNA Polymerase II
SDF-1a	stromal cell-derived factor-1 alpha
a-SNAP	soluble NSF attachment protein
SNAP-25	synaptosomal-associated protein of 25 kDa
SNARE	SNAP (soluble NSF attachment protein) receptor
SOCS	suppressor of cytokine signalling
SOE1	start of exon 1 (In FIE2, taken loosely to mean the TSS)
STAT	signal transducer and activator of transcription
TAF	TBP-associated factors
TBP	TATA-box binding protein
TF	transcription factor
TFBS	transcription factor binding site
TGN	trans-Golgi network
TMD	transmembrane domain
TNFa	tumor necrosis factor a
tPA	tissue plasminogen activator

TPR	tetratricopeptide repeat
TIS	translation initiation site
TSS	transcription start site
VAMP	vesicle-associated membrane protein
XBR	X2 box repressor