ISOLATION OF THE UDP GLUCURONOSYLTRANSFERASE 1A3 AND 1A4 PROXIMAL PROMOTERS AND CHARACTERIZATION OF THEIR DEPENDENCE ON THE TRANSCRIPTION FACTOR HEPATOCYTE NUCLEAR FACTOR 1 ALPHA

Dione A. Gardner-Stephen and Peter I. Mackenzie

Department of Clinical Pharmacology, Flinders University School of Medicine, Flinders

Medical Centre, Bedford Park, South Australia 5042, Australia

DMD #12203

Running Title: Regulation of UGT1A3 and UGT1A4 by HNF1a

Address correspondence to: Peter I. Mackenzie, Department of Clinical Pharmacology,

Flinders Medical Centre, Bedford Park, South Australia 5042, Australia

Phone: +61 8 8204 5394

Fax: +61 8 8204 5114

E-mail: Peter.Mackenzie@flinders.edu.au

Words in Abstract: 242

Words in Introduction: 343

Words in Results/Discussion: 1261

Number of References: 17

The abbreviations used are: UGT, UDP-glucuronosyltransferase; HNF, hepatocyte nuclear factor; EMSA, electrophoretic mobility shift assay

ABSTRACT

The UGT1A3-1A5 genes are a highly related UDP-glucuronosyltransferase (UGT) cluster, exhibiting high levels of coding and regulatory region homology. However, the ensuing proteins have both differing substrate specificities and differing expression patterns. The expression profile of each enzyme also varies considerably from one individual to the next. Differences in UGT expression have been predicted to contribute to an individual's response to pharmaceuticals, and to predisposition towards cancer in the event of carcinogen exposure. Therefore, it is desirable to elucidate the mechanisms that drive the transcription of UGT genes and identify the factors responsible for their variable expression. To this end we have isolated the UGT1A3, UGT1A4 and UGT1A5 proximal promoters and begun to investigate the regulatory elements necessary for activity in vitro. We have established that the nucleotide sequence upstream of the UGT1A5 exon 1 is an ineffective promoter, correlating with the lack of substantial expression of this UGT in human tissues. In contrast, the UGT1A3 and UGT1A4 proximal promoters are both highly active in hepatic and colonic cell lines, with maximal activity being encoded by the proximal 500 bp. However, the UGT1A3 and UGT1A4 promoters exhibit low activity in the human embryonic kidney cell line HEK293, unless coexpressed with hepatocyte nuclear factor (HNF) 1α . Furthermore, mutation of the consensuslike HNF1-binding site in the UGT1A3 promoter abolishes promoter function in all cell types. This study suggests an important role for HNF1 α in the transcriptional regulation of the human UGT1A3 and UGT1A4 genes.

DMD #12203

Many small lipophilic molecules require conjugation with a polar moiety before they can be efficiently excreted from the body. Catalyzed by members of UDPthe glucuronosyltransferase (UGT) enzyme superfamily, glucuronidation is a predominant example of this metabolic process, and one that usually results in a loss of bioactivity of the parent compound. Therefore, the efficacy and/or toxicity of many xenobiotics (pharmaceuticals, environmental contaminants and dietary compounds) and the fate of numerous endogenous signaling molecules (such as steroid hormones and neurotransmitters) can be regulated through glucuronidation (Radominska-Pandya et al., 1999; Tukey and Strassburg, 2000). Differences in UGT expression have been predicted to contribute to an individual's response to pharmaceuticals, and to their predisposition towards developing certain cancers. Thus it is desirable to elucidate the mechanisms that drive the transcription of *UGT* genes and identify the factors responsible for their variable expression.

UGT1A3, UGT1A4 and UGT1A5 are a trio of highly related proteins encoded by the human *UGT1A* locus. They share greater than 90% identity in their primary amino acid sequences and more than 85% nucleotide sequence identity in their proximal promoters to one kilobase (Green and Tephly, 1998; Gong et al., 2001). However, despite these similarities, this gene cluster varies considerably in both substrate specificity and expression pattern. While *UGT1A3* and *UGT1A4* mRNA transcripts have been found in many tissues including liver, biliary tissue, colon, and breast (Strassburg et al., 1997; Strassburg et al., 1998a; Chouinard et al., 2006), *UGT1A5* has not been found to be expressed to any significant extent in any tissues; although highly variable (but very low) expression in liver and gastrointestinal tract has been recently reported (Finel et al., 2005). Furthermore, there are inherent differences in the expression of *UGT1A3* and *UGT1A4* is not (Strassburg et al., 1998b). Because the mechanisms that determine the expression levels and/or tissue specificity of the *UGT1A3-1A5* cluster are currently not well understood, we have isolated their respective proximal promoters and begun to investigate the regulatory elements necessary for their activity *in vitro*.

METHODS

Isolation of the UGT1A3 and UGT1A4 promoters

The proximal 3.3 and 3.4 kb of the UGT1A3 and UGT1A4 promoters respectively were amplified by nested PCR from NotI digested human genomic DNA. Briefly, PfuTurbo (Stratagene, CA) was used to simultaneously amplify both promoters using the primers GTTATCATTAAATAATAATCCT and GGCAGGGGAACCTGGAGTCCT, and the resulting PCR product used as template to specifically amplify each promoter in separate reactions. The second round PCR primers were AGCCATAAGCTTATTGGATACCAGTATTGCT and AGCCATAAGCTTCTCAGCAGAAGACACGGACA for the UGT1A3 promoter, or AGCCATGCTAGCATCTGATATCAGTAATGTG and AGCCATCTCGAGCTCAGCAGAAGCCACCGACA for UGT1A4. Both promoters were

cloned into pGL3-basic (Promega, WI) and their ends sequenced to confirm their identity.

Generation of UGT1A3 and UGT1A4 promoter deletion constructs and mutants

The pGL3-1A3-3.3k and pGL3-1A4-3.4k vectors were used as templates to clone therequired deletion fragments of each promoter. The antisense primer sequences for allUGT1A3andUGT1A4fragmentswereAGCCATCTCGAGCTCAGCAGAAGACACGGACAand

AGCCATCTCGAGCTCAGCAGAAGCCACCGACA respectively. The sense primers annealed to bases -2541 to -2523, -1539 to -1519, -507 to -487, -200 to -184, -150 to -130 or -130 to -113 of the UGT1A3 promoter, relative to the translation start site. The three longest UGT1A4 promoter sub-fragments were amplified with the same sense primers as UGT1A3, giving lengths of 2610, 1574 and 506 nucleotides. For constructs containing less than 500 bases of the UGT1A4 promoter, PCR primers annealing to UGT1A4 nucleotides -200 to -184, -150 to -130 or -130 to -113 were used. To generate the mutated UGT1A3-150bp construct, the *UGT1A3*-150bp promoter was re-amplified with the sense primer AGCCATGCTAGCGGCCAACGCTTCACTAGAGGA containing the desired mutations (in bold). All resulting PCR products were cloned into pGL3-basic and sequenced in full.

Isolation of the UGT1A5 promoter

The proximal 1.5 kb of the *UGT1A5* promoter was also amplified using two sequential *PfuTurbo* reactions. The first round of PCR was performed on bacterial artificial chromosome 1308M2 from the human library RPCI-11 (BACPAC Resources, CA) using primers GAGGTCTTTAGACCACTTAGTC and GCTCCACACAAGACCTATGTATGAT. The resulting products were used as template to amplify the *UGT1A5*-1550bp promoter using the same primers as for *UGT1A4*-1574bp. The 508bp and 150bp fragments of the *UGT1A5* promoter were also amplified using the corresponding *UGT1A4* primers defined above. All fragments were ligated into pGL3-basic and sequenced in full.

Transient transfection and luciferase reporter assay

HepG2, Caco-2 and HEK293 cells were obtained from the American Type Tissue Collection. All transfections were performed using Lipofectamine 2000 (Invitrogen) in 24-well plates seeded with 2×10^5 HepG2, 7.5×10^4 Caco-2 or 1×10^5 HEK293 cells the previous day. Either 0.5 µg of empty pGL3-basic or a reporter vector carrying *UGT1A3*, *UGT1A4* or *UGT1A5* promoter sequences was co-transfected with 0.25 µg of pCMX-HNF1 α (Mackenzie et al., 2005) or empty pCMX vector. pRL-null (0.025µg) was added to all transfections as an internal control for transfection efficiency. The reporter and expression plasmid concentrations were optimized to maximize promoter response without exceeding the recommended DNA:Lipofectamine ratio or inducing cytotoxicity (Gardner-Stephen and Mackenzie, 2005). After 48 hours, cells were lysed in passive lysis buffer (Promega) and analyzed for firefly and renilla luciferase activity using the Dual-Luciferase Reporter Assay System (Promega) and a TopCount luminescence and scintillation counter (Parkard, Australia). As a minimum, all triplicate transfections were performed twice in independent experiments.

Electrophoretic mobility-shift assay (EMSA)

HNF1α EMSA, super-shift and competition assays were performed with Caco-2 and HepG2 nuclear extracts as described in Gardner-Stephen and Mackenzie (2005). The DNA probes utilized were: *UGT1A3*-HNF1 (ATTAATGGTTAATAATTAACTAGAGG); *UGT1A4*-

HNF1(ATTAATGGGTAATAAGTAACTGGTGG);UGT1A3-HNF1mut(ATTAATGGCCAACGCTTCACTAGAGG) and the unrelated probe sequence FXR-
consensus (GATCTCAAGAGGTCATTGACCTTTTG). Underlined text indicates the
extent of the putative HNF1-binding sites in each probe and mutations are highlighted in bold.

RESULTS AND DISCUSSION

Relatively little is currently known about the transcriptional regulation of the *UGT1A3-1A5* genetic cluster. Accordingly, we have initiated a detailed analysis of the *UGT1A3-1A5* promoters and hereby describe a number of fundamental functional similarities and differences between these highly related sequences.

Basal activities of the UGT1A3, UGT1A4 and UGT1A5 proximal promoters

In both liver and colon-derived cell lines, the UGT1A3 and UGT1A4-130bp promoters had minimal activity, exhibiting less than 3-fold increases over basal reporter gene expression by the promoter-less pGL3-basic vector. However, inclusion of a further 20 nucleotides of either promoter substantially increased luciferase expression in both cell types (Figure 1A and B). Further increases in promoter activity could be obtained in either cell line by inclusion of up to 500 base pairs of the UGT1A3 or UGT1A4 promoters, with the exception of the UGT1A3 promoter in Caco-2 cells, which showed greatest activity at a length of 200bp (Figure 1B). The largest increases obtained in promoter activity for UGT1A3 were 53-fold in HepG2 and 40-fold in Caco-2 cells; while the UGT1A4 promoter had maximal activities of 21 and 15 times that of pGL3-basic in the same cell lines (Figure 1A and B). Increasing promoter length beyond 500 nucleotides resulted in reduced promoter activity for both UGT genes, although this phenomenon was more marked in HepG2 cells than Caco-2. It was also found that, for all promoter lengths, the UGT1A3 gene had greater activity in vitro than UGT1A4, regardless of host cell type. Much of the increased activity of the UGT1A3 promoter over that of UGT1A4 was found to require nucleotides -150 to -200, although the transcription factors mediating this effect remain unidentified.

Comparison of the *UGT1A5* promoter with the regulatory regions of *UGT1A3* and *UGT1A4* revealed that the former had the least activity. This was true for all three promoter lengths tested, in both HepG2 and Caco-2 cells (Figure 2A and B). The *UGT1A5* promoter also differed from *UGT1A3* and *UGT1A4* in that the shortest fragment tested, 150 bp, was the most active. Increasing the promoter length to 500 nucleotides decreased promoter function in both HepG2 and Caco-2 cells, a result in direct opposition to that obtained for *UGT1A3* and

UGT1A4 (Figure 1A and B). UGT1A5 expression has not been detected at a substantial level in any human tissue to date, and it has been suggested that this may be due to lack of a functional promoter (Tukey and Strassburg, 2001). These results support this hypothesis, and suggest that although the *UGT1A5* core promoter is sufficient for assembly of a pre-initiation complex, one or more crucial regulatory elements between -150 and -500 bp are missing, and/or that the *UGT1A5* promoter contains negative regulatory sequences not present in *UGT1A3* and *UGT1A4*.

HNF1 is required for basal activity of the UGT1A3 and UGT1A4 proximal promoters

For both UGT1A3 and UGT1A4, it was found that nucleotides -130 to -150 were important for basal activity in HepG2 and Caco2 cells. This region of the UGT1A4 promoter has previously been predicted to contain an HNF1-binding site (Tronche et al., 1997), based on the high identity (11 of 13 nucleotides) of this region with the HNF1-binding site consensus sequence GTTAATNATTAAC. Furthermore, the equivalent region of the UGT1A3 promoter contains a 100% match to the HNF1-binding site consensus. As to date, however, no experimental evidence has been presented to ascertain whether these sites are functional in the context of their promoters. To investigate the influence of HNF1 α on the activity of the UGT1A3 and UGT1A4 promoters, constructs containing 500 bp or less of each regulatory region were co-transfected with HNF1a into cells known to express HNF1 factors (HepG2 and Caco-2 cells) or a cell line devoid of HNF1 factors; namely HEK293 (Bernard et al., 1999; Gardner-Stephen and Mackenzie, 2005). In HEK293 cells, it was found that UGT1A3 and UGT1A4 promoters of sufficient length to include the putative HNF1-binding site were highly responsive to heterologous expression of HNF1a. Reporter gene expression under the control of the UGT1A3 or UGT1A4 promoters could be increased up to 22-fold over basal levels (Figure 2C). In contrast, vectors containing only the most proximal 130 bp of the UGT1A3 or UGT1A4 promoters were completely unresponsive to the presence of HNF1a.

In HepG2 or Caco-2 cells, co-transfections of the *UGT1A3* or *UGT1A4* promoters with HNF1 α resulted in little or no additional response (Figure 2A and B). Since HepG2 and Caco-2 cells express HNF1 factors (Kuo et al., 1990; Rey-Campos et al., 1991), we postulated that

DMD #12203

the endogenous levels of HNF1 α and/or HNF1 β in these cells were sufficient to support expression of the reporter gene from the *UGT1A3* and *UGT1A4* promoters *in vitro*. Therefore, we mutated the putative HNF1-binding site in the *UGT1A3*-150bp promoter to abolish any binding of HNF1 α . The functional result of this mutation was a loss of basal activity of the *UGT1A3*-150 bp promoter in both HepG2 and Caco-2 cells, and prevention of HNF1 α responsiveness in HEK293 cells. In all cases, the mutated *UGT1A3*-150bp promoter construct behaved in the same manner as the *UGT1A3*-130bp promoter that contains no recognized HNF1 α -binding site (Figure 2). In support of the above evidence that the *UGT1A3* and *UGT1A4* promoter HNF1 sites are functional, we were able demonstrate that HNF1 factors from Caco-2 and HepG2 nuclear extracts can bind these sequences in EMSA (Figure 3). Furthermore, the mutation used to abolish HNF1 α -responsiveness of the *UGT1A3*-150bp promoter also prevented binding of HNF1 factors to this region *in vitro*, while neither the mutated or unrelated (FXR) probes could interfere with HNF1 binding when added in 500fold excess (Figure 3).

The UGT1A3 and UGT1A4 proximal promoters differ in their HNF1 responses

During the course of this study, two notable differences between the *UGT1A3* and *UGT1A4* HNF1-responses were observed. Firstly, while none of the *UGT1A3* promoter constructs exhibited any activity in HEK293 cells in the absence of HNF1 α , *UGT1A4* promoters of 150 bp or longer could support a small degree of basal activity. This activity, which was 2 to 3-fold greater than the empty vector control, is presumably HNF1-independent (Figure 2C). The second observation was that while *UGT1A3* promoter activity could not be increased in HepG2 cells by over-expression of HNF1 α , *UGT1A4* promoter activity was increased up to 2.3-fold by excess HNF1 α for promoter fragments \geq 150 bp (Figure 2A). One possible explanation is that the perfect *UGT1A3* HNF1-binding element is fully occupied at physiological HNF1 concentrations, while the slightly flawed site of the *UGT1A4* promoter is less efficient at competing with the multitude of genomic sites for limited HNF1. Therefore, addition of excess HNF1 α into the system can only increase the occupancy rate of the *UGT1A4* HNF1-binding site. There is also likely a cell-type-specific component to this

DMD #12203

second difference between the promoters, as it was only observed in cells of hepatic origin (Figure 2).

Summary

In summary, we have shown that while the nucleotide sequence upstream of the UGT1A5 first exon is ineffective as a promoter, the UGT1A3 and UGT1A4 proximal promoters are transcriptionally active *in vitro*. We have also shown that HNF1 α can interact functionally with both the UGT1A3 and UGT1A4 promoters, and that the HNF1 site of the UGT1A3 gene is essential for transcriptional activity *in vitro*. In this regard, the UGT1A3 promoter appears to be most similar to UGT1A1 (Bernard et al., 1999), and distinct from that of the human UGT1A8-1A10 gene cluster (Gregory et al., 2003). Since HNF1 factors are present in cell types where UGT1A3 and UGT1A4 are expressed (Kuo et al., 1990; Rey-Campos et al., 1991), it would be reasonable to expect that HNF1 α is important in UGT1A3 and UGT1A4transcription *in vivo*.

REFERENCES

- Bernard P, Goudonnet H, Artur Y, Desvergne B and Wahli W (1999) Activation of the mouse TATA-less and human TATA-containing UDP-glucuronosyltransferase 1A1 promoters by hepatocyte nuclear factor 1. *Mol Pharmacol* 56:526-536.
- Chouinard S, Tessier M, Vernouillet G, Gauthier S, Labrie F, Barbier O and Belanger A (2006) Inactivation of the pure antiestrogen fulvestrant and other synthetic estrogen molecules by UDP-glucuronosyltransferase 1A enzymes expressed in breast tissue. *Mol Pharmacol* 69:908-920.
- Finel M, Li X, Gardner-Stephen D, Bratton S, Mackenzie PI and Radominska-Pandya A (2005) Human UDP-glucuronosyltransferase 1A5: identification, expression, and activity. J Pharmacol Exp Ther 315:1143-1149.
- Gardner-Stephen DA and Mackenzie PI (2005) Identification and characterisation of functional Hepatocyte Nuclear Factor 1 binding sites in UDP-glucuronosyltransferase genes. *Methods Enzymol* 400:22-46.
- Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS and Popescu NC (2001) Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* 11:357-368.
- Green MD and Tephly TR (1998) Glucuronidation of amine substrates by purified and expressed UDP-glucuronosyltransferase proteins. *Drug Metab Dispos* 26:860-867.
- Gregory PA, Gardner-Stephen DA, Lewinsky RH, Duncliffe KN and Mackenzie PI (2003) Cloning and characterization of the human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 gene promoters: differential regulation through an initiator-like region. J Biol Chem 278:36107-36114.
- Kuo CJ, Conley PB, Hsieh CL, Francke U and Crabtree GR (1990) Molecular cloning, functional expression, and chromosomal localization of mouse hepatocyte nuclear factor 1. Proc Natl Acad Sci USA 87:9838-9842.

- Mackenzie PI, Gregory PA, Lewinsky RH, Yasmin SN, Height T, McKinnon RA and Gardner-Stephen DA (2005) Polymorphic variations in the expression of the chemical detoxifying UDP glucuronosyltransferases. *Toxicol Appl Pharmacol*. 207:S77-S83.
- Radominska-Pandya A, Czernik PJ, Little JM, Battaglia E and Mackenzie PI (1999) Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* 31:817-899.
- Rey-Campos J, Chouard T, Yaniv M and Cereghini S (1991) vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. *EMBO (Eur Mol Biol Organ) J* 10:1445-1457.
- Strassburg CP, Manns MP and Tukey RH (1998a) Expression of the UDPglucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8. J Biol Chem 273:8719-8726.
- Strassburg CP, Nguyen N, Manns MP and Tukey RH (1998b) Polymorphic expression of the UDP-glucuronosyltransferase UGT1A gene locus in human gastric epithelium. *Mol Pharmacol* 54:647-654.
- Strassburg CP, Oldhafer K, Manns MP and Tukey RH (1997) Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. *Mol Pharm* 52:212-220.
- Tronche F, Ringeisen F, Blumenfeld M, Yaniv M and Pontoglio M (1997) Analysis of the distribution of binding sites for a tissue-specific transcription factor in the vertebrate genome. J Mol Biol 266:231-245.
- Tukey RH and Strassburg CP (2000) Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Ann Rev Pharmacol Toxicol* 40:581-616.
- Tukey RH and Strassburg CP (2001) Genetic multiplicity of the human UDPglucuronosyltransferases and regulation in the gastrointestinal tract. *Mol Pharmacol* 59:405-414.

DMD #12203

FOOTNOTES

This work was supported by grants from the Cancer Council of South Australia and the National Health and Medical Research Council of Australia. P.I.M is a National Health and Medical Research Council Senior Principal Research Fellow.

Reprints: Peter I. Mackenzie, Department of Clinical Pharmacology, Flinders Medical Centre,

Bedford Park, South Australia 5042, Australia

E-mail: Peter.Mackenzie@flinders.edu.au

LEGENDS FOR FIGURES

Figure 1: Successive deletion constructs of the *UGT1A3* and *UGT1A4* promoters reveal positive regulatory elements necessary for basal activity.

A. HepG2 or B. Caco-2 cells were transfected in triplicate with 0.5 μ g of pGL3 reporter vectors carrying the indicated lengths of the *UGT1A3*, *UGT1A4* or *UGT1A5* promoters and 25 ng of the promoter-less control vector pRL-Null. Forty eight hours post-transfection the cells were lysed and assayed for firefly and *Renilla* luciferase reporter gene activities as described in "Methods". Representative results of at least two independent experiments are presented as the mean firefly:*Renilla* luciferase ratio relative to pGL3-basic (set arbitrarily to 1) ± standard deviation.

Figure 2: HNF1α is required for maximal basal activity of the *UGT1A3* and *UGT1A4* proximal promoters.

Mutations were introduced into the HNF1-binding site of the *UGT1A3*-150bp promoter by PCR as described in "Methods". A. HepG2, B. Caco-2 or C. HEK293 cells were co-transfected with 0.5 μ g pGL3-based vectors containing 500, 200, 150 or 130 nucleotides of the *UGT1A3* or *UGT1A4* promoters, 25 ng pRL-Null and 0.25 μ g of pCMX-HNF1 α expression vector. The DNA concentration in control transfections was kept constant by addition of empty pCMX vector as necessary. The results of all experiments are the means of triplicate samples, expressed as a relative value of firefly luciferase activity to the internal *Renilla* control, compared to the pGL3-basic control (set to 1). The error bars indicate one standard deviation.

Figure 3: HNF1 α binds to nucleotides -156 to -128 of the UGT1A3 and UGT1A4 promoters

Electrophoretic mobility-shift assays were performed using 50,000 cpm of ^{32}P end-labelled oligonucleotide probes encompassing the putative *UGT1A3* or *UGT1A4* HNF1 sites. HNF1 complexes were competed with 500-fold excess cold probe, or super-shifted with 2 µg HNF1a-specific antibody (Santa Cruz Biotechnology, CA). The positions of free probe and

DMD #12203

complexes containing HNF1a are indicated by arrows. WT: UGT1A3 or UGT1A4 wild-type

probe; Mut: UGT1A3 mutant probe; FXR: consensus FXR-binding sequence probe.

Figure 1

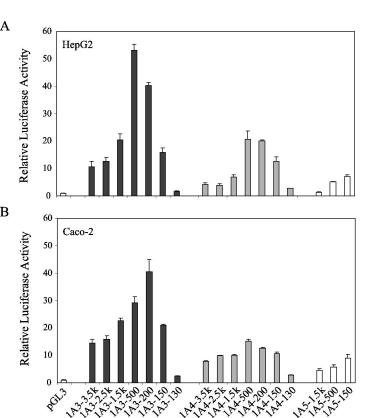
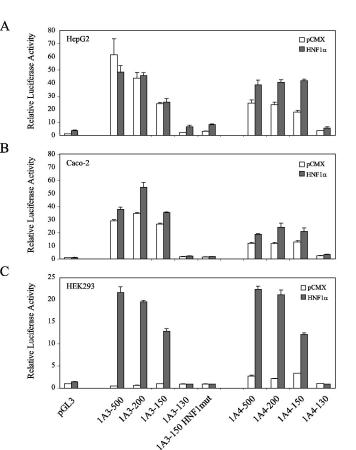


Figure 2



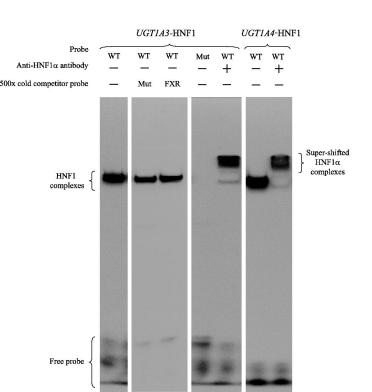


Figure 3