TOXICOLOGICAL SCIENCES **135(2)**, 390–401 2013
doi:10.1093/toxsci/kft166
Advance Access publication July 27, 2013

Bioinformatic Analysis of 302 Reactive Metabolite Target Proteins. Which Ones Are Important for Cell Death?

Robert P. Hanzlik,*[,1] Yakov M. Koen,* and Jianwen Fang†

*Department of Medicinal Chemistry and †Applied Bioinformatics Lab, University of Kansas, Lawrence, Kansas 66045

[1]To whom correspondence should be addressed at Department of Medicinal Chemistry, University of Kansas, 1251 Wescoe Hall Drive, Room 4048, Malott Hall, Lawrence, KS 66045. E-mail: rhanzlik@ku.edu.

Received February 18, 2013; accepted July 17, 2013

Many low molecular weight compounds undergo biotransformation to chemically reactive metabolites (CRMs) that covalently modify cellular proteins. However, the mechanisms by which this covalent binding leads to cytotoxicity are not understood. Prior analyses of lists of target proteins sorted by functional categories or hit frequency have not proven informative. In an attempt to move beyond covalent binding, we hypothesized that xenobiotic posttranslational modification of proteins might disrupt important protein-protein interactions (PPIs) and thereby direct cells from homeostasis into cell death pathways. To test this hypothesis, we analyzed a list of 302 proteins (66% rat, 26% mouse, 5% human) known to be targeted by 41 different cytotoxic CRMs. Human orthologs of rodent proteins were found by blast sequence alignment, and their interacting partners were found using the Human Protein Reference Database. The combined set of target orthologs and partners was sorted into KEGG pathways and Gene Ontology categories. Those most highly ranked based on sorting statistics and toxicological relevance were heavily involved with intracellular signaling pathways, protein folding, unfolded protein response, and regulation of apoptosis. Detailed examination revealed that many of the categories were flagged primarily by partner proteins rather than target proteins and that a majority of these partners interacted with just a small number of proteins in the CRM target set. A similar analysis performed without the partner proteins flagged very few categories as significant. These results support the hypothesis that disruption of important PPIs may be a major mechanism contributing to CRM-induced acute cytotoxicity.

Key Words: **reactive metabolites; protein covalent binding; protein-protein interaction; cytotoxicity.**

Posttranslational modification (PTM) is a well-known mechanism for regulating the activity of cellular proteins and the metabolic and signaling pathways they comprise (Krishna and Wold, 1998; Reinders and Sickmann, 2007). In contrast, xenobiotic PTMs stemming from reactive chemicals or chemically reactive metabolites (CRMs) are frequently associated with toxic consequences for the cell (Evans et al., 2004; Liebler, 2006; Liebler and Guengerich, 2005; Park et al., 2005). The enzymes that convert relatively innocuous chemicals into CRMs, and the covalent modification of cellular proteins by the latter, have been studied intensively for more than 40 years (Hanzlik et al., 2009; Park et al., 2011; Snyder, 2011). Early interest focused primarily on the small molecule precursors, the enzymes that bioactivated them, and the structures and reactivity of the CRMs formed. Later, improvements in protein separation technology facilitated the isolation and identification of individual CRM target proteins. Initially, the quest to identify target protein identification was driven by the hope that knowing these targets would provide mechanistic understanding of CRM cytotoxicity. Ironically, many of the early identified targets were enzymes (e.g., aldehyde dehydrogenase, cytochromes P450 2B1, 2C11, and 2E1, liver carboxylesterase, dipeptidyl peptidase IV, alpha-ketoglutarate dehydrogenase, and others), but none appeared to be critical for cell survival in the context of acute cytotoxicity (Hanzlik et al., 2009).

As increasing numbers of target proteins became known, it was necessary to take a broader view and not look just for a few potentially "critical" targets. For example, grouping targets according to their function, as shown in Table 1 for 62 targets of thiobenzamide metabolites (Ikehata et al., 2008), showed that some categories were more highly populated than others, but the categories were too broad to reveal mechanisms or even suggest plausible mechanistic hypotheses. Another approach to identifying potentially important target proteins is to rank them according to how frequently they are reported as targets of various CRMs. Such a compilation is listed in Table 2. This listing again contains some enzymes but also a number of structural proteins (e.g., albumin, hemoglobin, actin, tropomyosin, and tubulin) that would not immediately appear to be important for direct acute cytotoxicity. Finally, bioinformatic approaches such as Ingenuity Pathway Analysis (Druckova et al., 2007) and similar *ad hoc* approaches based on sorting target proteins according to KEGG pathways or the Gene Ontology (GO) classification schemes (Fang et al., 2009)

TABLE 1
Functional Classification of Cytosolic and Microsomal Protein Targets of Thiobenzamide Metabolites

Functional class	No. of cyt[a]	No. of mic[a]
Binding/Carrier proteins	6	7
Cytoskeleton/Structural proteins	1	1
Receptors/Signal transduction	0	1
Intermediary metabolism	18	0
Xenobiotic metabolism	2	3
Redox regulation	5	2
Protein folding, heat shock, and stress response	1	10
Protein degradation	2	0
Nucleic acid metabolism	3	0
Other	1	1

Note. [a]Number of cytosolic (cyt) and microsomal (mic) target proteins found in each functional class. The data are compiled from a listing by Ikehata et al. (2008).

TABLE 2
Reactive Metabolite Target Proteins Ranked by Number of CRMs That Target them[a]

Entry	No. of hits	Target protein name(s) (rat, mouse)
1	12	Protein disulfide isomerase, PDI, Erp59
2	12	Protein disulfide isomerase, PDI A3, ER-60
3	9	Heat shock 60 KD protein 1 (chaperonin), HSP-60
4	9	Albumin
5	8	Selenium-binding protein 2, 56 kDa acetaminophen-binding protein
6	8	Tropomyosin 3, tropomyosin gamma, tropomyosin alpha-3 chain
8	7	GAPDH, glyceraldehyde-3-phosphate dehydrogenase
9	7	Hemoglobin beta chain
10	7	Mortalin, GRP 75
11	7	Triosephosphate isomerase, TIM
12	6	Actin, cytoplasmic 2; gamma-actin
13	6	Alpha-enolase, Enolase 1
14	6	Alpha-tubulin
15	6	Heat shock 70 kDa protein 8, Hsc 70
16	6	Peroxiredoxin 6
17	5	Aldehyde dehydrogenase, mitochondrial
18	5	ATP synthase subunit beta
19	5	Carbonic anhydrase 3, CA-III
20	5	Beta-diketonase, Fumarylacetoacetate hydrolase
21	5	Hemoglobin alpha chain
22	4	Alpha-1 protease inhibitor 1, Alpha-1-antitrypsin 1-1
23	4	ATP synthase subunit alpha (mitochondrial)
24	4	Calreticulin
25	4	Cytochrome P450 2B1
26	4	Elongation factor 1-alpha 1
27	4	Fatty acid–binding protein 1

Note. [a]Data taken from reactive metabolite target protein database, http://tpdb.medchem.ku.edu:8080/protein_database/

have also been tried, but these approaches too have not been particularly helpful in illuminating mechanisms of toxicity or deciding which target proteins may be more important than others. Thus, the hope of understanding mechanisms of cytotoxicity by analyzing target proteins remains unfulfilled.

In this article, we report the application of a new approach to winnowing the large number of known target proteins in the Reactive Metabolite Target Protein Database (TPDB, 2013) to find those whose covalent modification by CRMs is most strongly related to cytotoxic outcomes. It is based on the hypothesis (Fang et al., 2009) that the detrimental consequences of protein modification by CRMs may stem not from the partial loss of enzymatic activities, but rather from the disruption or inadvertent mimicry of important protein-protein interactions (PPIs). This new approach, therefore, involves consideration of not just target proteins but also of their immediate partner proteins. Using this approach leads to a "ranking of target proteins" significantly different from that produced by considering target proteins alone (Table 2). We suggest that ranking target proteins this way may provide a better perspective on their relevance to toxicity and a better guide for formulating testable mechanistic hypotheses about their potential role in CRM cytotoxicity.

MATERIALS AND METHODS

The basic method of our analysis is outlined schematically in Figure 1. As in several previous analyses, the KEGG (www.genome.jp/kegg/) and GO (www.geneontology.org/) classification schemes were used to sort proteins objectively and obtain a statistical sense of the extent to which any particular pathway or category is significantly enriched in proteins from the list being sorted. A key difference in this analysis is that proteins reported to interact directly (i.e., physically) with any of the CRM target proteins are also included in the overall sorting operation. The partner proteins were included for two reasons. First, earlier analyses based solely on target proteins were not very fruitful. Second, the partner proteins are needed to represent fully the PPIs in which the target proteins participate.

For this analysis, a list of 302 proteins known to be adducted by one or more of 41 different CRMs was taken from the Reactive Metabolite Target Protein Database (TPDB, 2013). Sixty-six percent of all proteins in the TPDB were from rat, 26% were from mouse, and 5% were from human. By tissue, the distribution of target proteins was 57% liver, 16% lung, 13% kidney, 8% nervous, and 3% blood cells. Although 27 of these proteins were targeted by ≥ 4 different CRMs (Table 2), a great many were reported as targets of only one or two CRMs (TPDB "rank by hits" function). Considering only those proteins targeted by two or more CRMs reduced the total number of targets to 70. From these, we eliminated hemoglobin and one EST of unknown function giving a total of 68 target proteins of interest, mostly from rat or mouse or in some cases both. To eliminate interspecies redundancy and enable searching for partner proteins, human orthologs of the rat and mouse proteins were found by searching target proteins against human proteins using BLAST sequence alignment (Hanzlik et al., 2009). This gave a list of 52 nonredundant human orthologs (Table 3) whose interacting partners were identified using the Human Protein Reference Database (www.hprd.org). Of the 52 human protein orthologs, 16 had no reported partner proteins, whereas 36 others had a total of 327 known partners. A spreadsheet was made listing all the partners for each target (Supplementary table S1), and a copy was then re-sorted (inverted) to list the target protein(s) that invoked each individual partner protein (Supplementary table S2). The ability to look up which target invoked a particular partner protein, and conversely, becomes important later in the analysis.

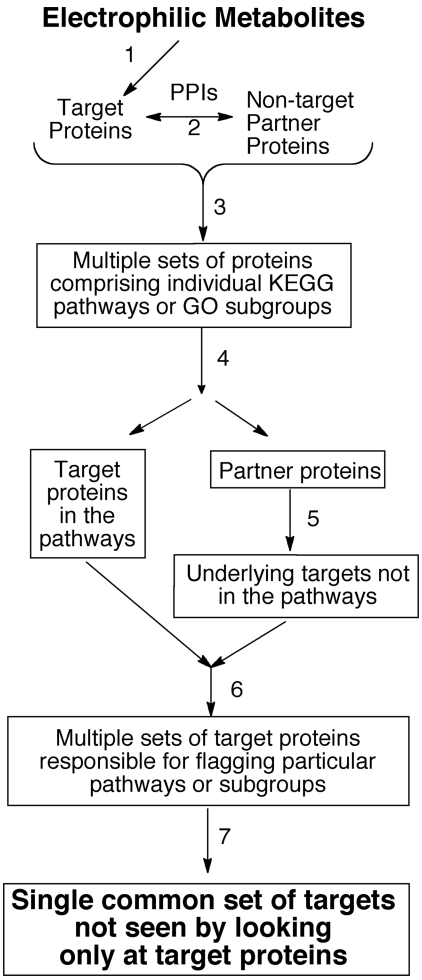

FIG. 1. Strategy for identification of important target proteins among others. The strategy is based on the idea that covalent modification of proteins by electrophilic metabolites (step 1) mimics or interferes with PPIs. After identification of a large number of target proteins, their interacting partners are found through PPI databases such as the HPRD (step 2). Target proteins and their partners are then combined, and the overall set is sorted into bins corresponding to KEGG pathways or GO categories (step 3). This yields a large number of bins that are evaluated for statistical significance and likely relevance to cytotoxicity. Some bins contain many partners and few or no targets. For each bin deemed significant and relevant, the proteins that sorted into it are then parsed (step 4) into a target protein set or a partner protein set. Step 5 is a key step. In this step, the residual partner proteins are then analyzed to determine which target proteins invoked them in the first place. For each bin, the underlying targets are combined with the targets found in step 4. The result is many sets of target proteins containing much redundancy among them. This redundancy gives insight into the ability of a given target protein to influence, through PPIs, multiple pathways of importance to the cell. Compressing the redundancy in the lists into numbers of occurrences for each protein (step 7) gives a list of target proteins, rank ordered by likely importance to toxicity, as judged by the number of PPIs that may be affected by their covalent modification. In this way, a list of 302 proteins in the TPDB was reduced to a short list; the top 10 or 15 of which have a very broad "reach" into cellular pathways and may therefore be more important than many other known target proteins.

Once all the known partner proteins were found, the 52 targets and 327 partners were combined into a single set of 379 proteins. This set was then sorted into "bins" in the form of GO categories and KEGG pathways using The Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2009). To identify bins of potential interest in relation to mechanisms of cytotoxicity, we initially used low stringency sorting criteria, namely, that there be at least two target and/or partner proteins in the bin and that $p < 0.1$. Because of the hierarchical nature of the classification schemes, the total number of bins found was much greater than the total number of proteins (see Results section and Supplementary tables S3, S4, and S5). To reduce the total number of pathways and categories populated, we examined the statistics of their filling including count, p value, and fold enrichment. Count gives the number of proteins that sorted into any given bin. A high count might indicate greater biological relevance of that bin (and the proteins in it) to CRM-induced cytotoxicity. A very low p value indicates the improbability of the proteins in a bin arriving there by chance. Finally, the fold enrichment value for a particular bin is the ratio of the fraction of the sorted set appearing in that bin versus the fraction of all classified proteins that sorts into that bin. A high fold enrichment could suggest a higher relevance of that bin and its proteins to CRM-induced cytotoxicity. To simplify the consideration of these statistical parameters, we created a composite index (CI), defined as CI = (fold enrichment) × (−log p), and used it to rank the various KEGG pathways and GO categories. The full rank ordered but unedited lists of all categories (bins) from the KEGG and GO sorting operations can be found in Supplementary tables S3, S4, and S5.

The ranking process ordered the categories but did not reduce their number, which remained much larger than the number of proteins entered to start with (see Results section). Furthermore, the quantitative indices of category rank formed a more or less continual distribution with no pronounced gaps or breaks. Therefore, to reduce the redundancy and shift the focus toward the more important proteins, the final selection of categories for further detailed analysis was based on their likely relevance to cytotoxicity. This step was not undertaken lightly because it runs the risk of introducing bias. Nevertheless, we found impossible to proceed further based purely on mathematical or statistical considerations. We therefore excluded pathways and categories that seemed unlikely to be related to or contribute to acute cytotoxicity. For example, we eliminated pathways and categories devoted to metabolism of plant natural products, bacterial virulence factors, central nervous system disorders, Alzheimer's disease, and cancer, as these were judged to be of little or no direct relevance for acute CRM-induced cytotoxicity in mammalian cells. We also excluded pathways related to blood clotting because CRM-induced cytotoxicity is well known to occur under cell culture conditions in the absence of blood and clotting factors. Finally, we eliminated categories related to metabolism of DNA, RNA, and complex structural lipids, as these processes generally occur on time scales much slower than acute cytotoxic responses to CRMs. In this way, we were able to come up with a more reasonable subset of flagged bins and proceed with their analysis toward the objective of determining which *target* proteins were most strongly connected to cytotoxic outcomes.

In addition to the statistical parameters associated with each bin, the sorting program DAVID also furnished a list of the gene names of all the target and partner proteins that had sorted into each bin. Using the spreadsheet lists of targets versus partners and partners versus targets mentioned above (Supplementary tables S1 and S2), it was easy to subdivide the proteins in each bin into a target subset and a partner subset, and for each partner protein to find the target protein(s) that originally invoked it. This last step was vital as some bins were found to contain few or even no known target proteins (see Tables 4 and 5 and Supplementary table S3 for examples). Although most target proteins have several or more partner proteins, only a few pairs of target proteins interact directly with each other. These observations emphasize the value and importance of including partner proteins in the analysis. It may also explain why earlier analyses using only target proteins gave no strong clues as to mechanisms or pathways of cytotoxicity and why they failed to distinguish potentially important targets from other targets. Finally, by analyzing a large number of KEGG and GO bins selected for high fold enrichment, low p value, and likely relevance to cytotoxicity, it became clear that a relatively small number of target proteins were consistently responsible for invoking a majority of the partner proteins that sorted into bins relevant for cytotoxicity.

TABLE 3
Human Orthologs of CRM Target Proteins Listed in the Reactive Metabolite Target Protein Database

Entry	Protein name	Gene name	Target proteins having known partners (n = 36)
1	P63261	ACTG1	Actin, cytoplasmic 2; gamma-actin
2	P00326	ADH1C	Alcohol dehydrogenase 1C
3	P42330	AKR1C3	Aldo keto reductase family 1, member C3
4	P02768	ALB	Albumin
5	P05091	ALDH2	Aldehyde dehydrogenase, mitochondrial
6	P05062	ALDOB	Aldolase B, fructose-bisphosphate
7	P05089	ARG1	Arginase
8	A5A6H5	ATP5A1	ATP synthase, alpha
9	P06576	ATP5B	ATP synthase, beta
10	P07451	CA3	Carbonic anhydrase III
11	P27797	CALR	Calreticulin
12	P04040	CAT	Catalase
13	P27487	DPP4	Dipeptidyl peptidase IV
14	P30040	ERP29	Endoplasmic reticulum protein 29
15	Q05CP7	FABP1	Fatty acid–binding protein 1
16	P08263	GSTA1	Glutathione S-transferase alpha 1
17	P09488	GSTM1	Glutathione S-transferase Mu-1
18	P09211	GSTP1	Glutathione S-transferase Pi
19	Q6PK50	HSP90AB1	HSP90B, HSP 84
20	Q96GW1	HSP90B1	GRP94, tumor rejection antigen 1
21	P11021	HSPA5	BIP, GRP 78, HSPA5
22	P11142	HSPA8	Heat shock 70 kDa protein 8, Hsc 70
23	P10809	HSPD1	Heat shock 60 KD protein 1 (chaperonin), HSP-60
24	O75874	IDH1	Isocitrate dehydrogenase 1
25	P07237	P4HB	Protein disulfide isomerase A1, PDI, Erp59
26	P30101	PDIA3	Protein disulfide isomerase A3, PDI A3, ER-60
27	Q15084	PDIA6	Protein disulfide isomerase A6
28	P30086	PEBP1	Phosphatidylethanolamine-binding protein 1, Raf kinase inhibitor protein
29	P00558	PGK1	Phosphoglycerate kinase 1
30	P32119	PRDX2	Peroxiredoxin 2
31	P30041	PRDX6	Peroxiredoxin 6
32	Q13228	SELENBP1	Selenium-binding protein 1
33	P01009	SERPINA1	Protease inhibitor 1, Alpha-1-antitrypsin 1-1
34	P09493	TPM1	Tropomyosin alpha 1 chain
35	P02766	TTR	Transthyretin
36	P10599	TXN	Thioredoxin
			Target proteins having no reported partners (n = 16)
1	P49189	ALDH9A1	4-trimethylaminobutyraldehyde dehydrogenase
2	P60709	ACTB	Actin, cytoplasmic 1; beta-actin
3	P06733	enol	Alpha-enolase
4	Q9UBR1	UPB1	Beta-alanine synthase, ureidopropionase
5	Q8TDZ9	CEShBr2	Brain carboxylesterase hBr2
6	P20813	CYP2B6	Cytochrome P450 2B6
7	P30046	ddt	D-dopachrome decarboxylase
8	P68104	EEF1A1	Elongation factor 1-alpha 1
9	P07099	ephX1	Epoxide hydrolase 1
10	P16930	FAH	Beta-diketonase, Fumarylacetoacetase
11	Q14353	Gamt	Guanidinoacetate N-methyltransferase
12	P28330	acadl	Long-chain-specific acyl-CoA dehydrogenase, mitochondrial
13	Q02252	Aldh6a1	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial
14	P52758	Hrsp12	14.5 kDa translational inhibitor protein, Ribonuclease UK114
15	Q00266	mat1a	S-adenosylmethionine synthase isoform type-1
16	Q53HE2	TPI1	Triosephosphate isomerase, TIM

RESULTS AND DISCUSSION

Our objective was to analyze a list of CRM target proteins to identify those whose modification by CRMs was most strongly linked to cytotoxicity or in other words to identify which targets among the many were potentially most important to causing cytotoxicity upon modification by CRMs. Directly interacting partner proteins were included in the analysis to represent the presumably important PPIs of the chemically modified target proteins. Sorting the 379 target and partner proteins into KEGG and GO gave the following results: 199 proteins sorted into 46 KEGG pathways; 345 proteins sorted into 466 GO Biological Process (BP) categories; 355 proteins sorted into 119 GO Molecular Function (MF) categories; and 343 proteins sorted into 85 GO Cellular Component categories. We did not analyze the cellular component category further because in many cases target proteins had been identified after subcellular fractionation. Instead, we performed a detailed analysis of those target and partner proteins that sorted into the KEGG, GO-BP, and GO-MF classification schemes. A key part of the analysis was to use partner proteins to highlight the CRM target proteins that invoked them in the first place (see Materials and Methods). Once the three main sortings were the analyzed and the results rank ordered, the results were pooled into a single list containing only target proteins rank ordered for likely relevance to acute cytotoxicity. The analyses are described in the following sections.

TABLE 4
Ten Top-Ranking KEGG Pathways and Their Statistical Parameters

Subcategory groups	Count	*p* value	Fold enrichment	CI[a]
Group 1. Signaling pathways				
hsa04620:Toll-like receptor signaling pathway	20	1.59E–07	4.15	28.23
hsa04920:Adipocytokine signaling pathway	13	1.21E–04	3.77	14.77
hsa04010:MAPK signaling pathway	26	3.75E–04	2.13	7.28
hsa04012:ErbB signaling pathway	11	6.00E–03	2.74	6.09
hsa04310:Wnt signaling pathway	14	2.40E–02	1.96	3.18
hsa04370:VEGF signaling pathway	8	4.76E–02	2.39	3.16
Group 2. Antigen processing				
hsa04612:Antigen processing and presentation	14	7.14E–05	3.71	15.37
Group 3. Apoptosis related				
hsa04210:Apoptosis	13	4.73E–04	3.28	10.90
Group 4. Xenobiotic metabolism	6	3.42E–02	3.26	4.78
hsa00480:Glutathione metabolism	7	9.82E–02	2.18	2.20
hsa00980:Metabolism of xenobiotics by cyt P450	18	3.61E–07	4.38	28.23

Note. For the gene names associated with each subcategory, see Table 5.
[a]Composite index defined as (fold enrichment) × (−log p).

TABLE 5
Gene Names of Proteins in Ten Top-Ranked KEGG Pathways Listed in Table 4

hsa04620: Toll-like receptor signaling pathway	hsa04920: Adipocytokine signaling pathway	hsa04010: MAPK signaling pathway	hsa04012: ErbB signaling pathway	hsa04310: Wnt signaling pathway	hsa04370: VEGF signaling pathway	hsa04612: Antigen processing and presentation	hsa04210: Apoptosis	hsa00480: Glutathione metabolism	hsa00980: Metabolism of xenobiotics by cyt P450
AKT1	AKT1	AKT1	ABL2	CSNK2A1	AKT1	B2M	AKT1	***GSTA1***	***ADH1C***
CCL5	CHUK	ATF4	AKT1	CSNK2A2	CASP9	***CALR***	CASP3	GSTA2	***AKR1C3***
CD40	IKBKB	CASP3	ERBB2	JUN	MAP2K1	CANX	CASP7	***GSTM1***	***GSTA1***
CHUK	IKBKG	CHUK	JUN	MAP3K7	MAPK1	CD4	CASP9	GSTM2	GSTA2
CXCL10	JAK2	IKBKB	MAP2K1	MAPK8	PRKCA	HLA-A	CHUK	***GSTP1***	***GSTM1***
CXCL11	MAPK8	IKBKG	MAPK1	MMP7	PRKCB1	HLA-C	CYCS	***IDH1***	GSTM2
CXCL9	NFKB1	JUN	MAPK8	PPARD	PRKCG	HSP90AA1	IKBKB	6	***GSTP1***
IKBKB	PPARA	MAP2K1	PRKCA	PPP2R1A	RAF1	***HSP90AB1***	IKBKG		7
IKBKE	PTPN11	MAP3K14	PRKCB1	PPP2R1B	8	***HSPA5***	MAP3K14		
IKBKG	SLC2A1	MAP3K3	PRKCG	PRKACA		***HSPA8***	NFKB1		
JUN	SLC2A4	MAP3K5	RAF1	PRKCA		***PDIA3***	PRKACA		
MAP2K1	STAT3	MAP3K7	11	PRKCB1		PSME3	TP53		
MAP3K7	TRAF2	MAPK1		PRKCG		TAP1	TRAF2		
MAPK1	13	MAPK8		TP53		TAPBP	13		
MAPK8		MAPT		14		14			
NFKB1		NFKB1							
TBK1		PDGFRB							
TLR1		PRKACA							
TLR2		PRKCA							
TLR4		PRKCB1							
20		PRKCG							
		RAF1							
		RASA1							
		STMN1							
		TP53							
		TRAF2							
		26							

Note. The gene names in boldface italics correspond to target proteins; all others correspond to partner proteins. The numbers at the bottom of each column are the count values for the column; the total for all 10 columns is 132.

KEGG Pathway Analysis of Targets and Partners

After rank ordering the 46 KEGG pathway bins by their CI value (Supplementary table S3), 10 were selected for further detailed analysis. They included six intracellular signaling pathways, two pathways related to metabolism of xenobiotics, and one each related to antigen processing and presentation and apoptosis (Table 4). The sorting process also yielded lists of genes representing each of the proteins in each of the bins. As an example, the 132 genes associated with the 10 pathways in Table 4 are listed in Table 5. Several of these genes appear in more than one pathway, but a more striking observation is that among the six intracellular signaling pathways, *none* of the genes correspond to actual CRM target proteins. A likely reason for this is that although the target proteins have definite cellular functions, they are not actually members of the signaling pathways, even though they may interact with one or more of the pathway members. The converse issue of why the signaling pathway members are not detected as CRM targets is almost certainly related to analytical sensitivity, and the fact that the cellular concentrations of these signaling proteins are typically very low compared with target proteins detectable by the usual proteomic methods (Koen et al., 2000). In contrast, a majority of the genes in the pathways related to xenobiotic metabolism and antigen processing and presentation do correspond directly to target proteins (Table 5). The absence of genes for CRM targets among the signaling pathway genes

TABLE 6
Example of Spreadsheet Used for Manual Counting of Target Proteins Associated With Partner Proteins in KEGG Pathways Such As Those in Table 5

A. Gene names of target proteins whose partners are listed in the KEGG pathways in Tables 4 and 5				B. Gene names of common denominator target proteins (redundancies removed from part A)				
TLR	MAPK	Antigen processing	Apoptosis	TLR	MAPK	Antigen processing	Apoptosis	Number of occurrences
DPP4	GSTM1	ATP5B	GSTP1			ATP5B		1
DPP4	GSTP1	CALR	HSP90AB1			CALR		1
DPP4	GSTP1	CALR	HSP90AB1	DPP4		DPP4		2
DPP4	HSP90AB1	CALR	HSP90AB1		GSTM1			1
GSTP1	HSP90AB1	CALR	HSP90AB1	GSTP1	GSTP1		GSTP1	**3**
HSP90AB1	HSP90AB1	DPP4	HSP90AB1	HSP90AB1	HSP90AB1	HSP90AB1	HSP90AB1	**4**
HSP90AB1	HSP90AB1	HSP90AB1	HSP90AB1	HSP90B1				1
HSP90AB1	HSP90AB1	HSPA5	HSPA5		HSPA5	HSPA5	HSPA5	**3**
HSP90AB1	HSP90AB1	HSPA5	HSPA8	HSPA8	HSPA8	HSPA8	HSPA8	**4**
HSP90AB1	HSP90AB1	HSPA5	HSPD1	HSPD1	HSPD1		HSPD1	**3**
HSP90AB1	HSPA5	HSPA8	HSPD1			PDIA3		1
HSP90B1	HSPA8	HSPA8	HSPD1	PEBP1	PEBP1		PEBP1	**3**
HSP90B1	HSPA8	PDIA3	PEBP1		PRDX2			1
HSP90B1	HSPA8	PDIA3	PEBP1	TPM1	TPM1			2
HSPA8	HSPD1	PDIA3	PEBP1		TTR			1
HSPD1	HSPD1	PDIA3	TXN	TXN	TXN		TXN	**3**
PEBP1	HSPD1	**Total 17**	**Total 17**					
PEBP1	PEBP1							
PEBP1	PEBP1							
PEBP1	PEBP1							
TPM1	PEBP1							
TXN	PEBP1							
Total 22	PEBP1							
	PEBP1							
	PEBP1							
	PEBP1							
	PRDX2							
	TPM1							
	TTR							
	TXN							
	TXN							
	Total 30							

Note. The left-hand part of this table (A) shows the redundancy among targets in any given bin from Table 5. This redundancy arises from the fact that one target protein may have many partners that sort into different KEGG pathway bins. The total number of target proteins listed at the bottom of each column is slightly larger than the number of partner proteins shown in Table 5 because several partners are invoked by more than one target. After collapsing this redundancy, the right-hand half of the table (B) shows that certain target proteins may be responsible for invoking partners in several different pathways. This redundancy is collapsed by counting the number of pathways in which a target protein may influence partners, even though the target protein may not formally belong to that pathway. In this way, it can be seen that 73 of the nearly 115 target proteins listed in Table 5 are invoked by just the seven proteins highlighted in boldface.

TABLE 7
Top-Ranking GO-BP Subcategories (n = 46) and Their Statistical Parameters, Grouped as Described in the Text

GO subcategory group	Count	p value	Fold enrichment	CI
Group 1. Apoptosis, its regulation, and its effects				
GO:0051014~actin filament severing	4	4.42E−05	44.4	193.3
GO:0006915~apoptosis	66	1.67E−20	3.7	73.9
GO:0012501~programmed cell death	66	2.66E−20	3.7	72.5
GO:0042981~regulation of apoptosis	50	3.38E−17	4.1	68.0
GO:0016265~death	67	9.50E−20	3.6	67.8
GO:0008219~cell death	67	9.50E−20	3.6	67.8
GO:0043067~regulation of programmed cell death	50	5.29E−17	4.1	66.4
GO:0008632~apoptotic program	13	2.75E−07	7.0	46.2
GO:0043066~negative regulation of apoptosis	24	2.22E−09	4.6	40.1
GO:0043069~negative regulation of programmed cell death	24	2.86E−09	4.6	39.1
GO:0006916~antiapoptosis	19	4.84E−08	5.0	36.5
GO:0006917~induction of apoptosis	17	2.60E−05	3.5	16.2
GO:0012502~induction of programmed cell death	17	2.75E−05	3.5	16.1
GO:0043065~positive regulation of apoptosis	19	1.81E−05	3.3	15.7
GO:0043068~positive regulation of programmed cell death	19	2.01E−05	3.3	15.5
Group 2. ER stress, protein folding, and unfolded protein response				
GO:0065008~regulation of biological quality	75	1.02E−23	3.8	86.9
GO:0006983~ER overload response	3	4.81E−03	26.6	61.7
GO:0006457~protein folding	31	8.42E−13	5.0	60.2
GO:0030968~unfolded protein response	5	2.09E−04	15.9	58.3
GO:0006986~response to unfolded protein	15	2.20E−08	7.1	54.2
Group 3. NO and inflammation				
GO:0046209~nitric oxide metabolic process	7	1.94E−05	12.0	56.3
GO:0006809~nitric oxide biosynthetic process	7	1.94E−05	12.0	56.3
GO:0006954~inflammatory response	30	4.11E−11	4.4	46.0
Group 4. Oxidative stress				
GO:0006979~response to oxidative stress	11	5.84E−05	5.1	21.5
Group 5. Kinases and signaling cascades				
GO:0043549~regulation of kinase activity	20	1.65E−06	3.8	21.8
GO:0033673~negative regulation of kinase activity	8	3.05E−04	6.1	21.5
GO:0006469~negative regulation of protein kinase activity	8	3.05E−04	6.1	21.5
GO:0033674~positive regulation of kinase activity	13	6.57E−05	4.2	17.5
GO:0007243~protein kinase cascade	26	3.05E−06	2.9	16.2
GO:0043405~regulation of MAP kinase activity	10	3.56E−04	4.5	15.6
GO:0000165~MAPKKK cascade	13	3.18E−04	3.5	12.4
GO:0007242~intracellular signaling cascade	63	1.43E−06	1.9	10.9
GO:0007165~signal transduction	123	3.05E−06	1.5	8.0
GO:0007267~cell-cell signaling	30	4.67E−04	2.0	6.7
Group 6. Homeostasis, regulation, wounding, and general stress				
GO:0009611~response to wounding	39	5.24E−13	4.0	49.3
GO:0042592~homeostatic process	37	4.45E−12	3.9	44.6
GO:0033554~cellular response to stress	5	4.72E−04	13.1	43.4
GO:0006952~defense response	44	4.23E−12	3.4	38.5
GO:0048518~positive regulation of biological process	64	4.95E−12	2.6	29.0
GO:0048519~negative regulation of biological process	64	7.90E−11	2.4	24.3
GO:0048522~positive regulation of cellular process	56	4.44E−10	2.5	23.3
GO:0048523~negative regulation of cellular process	61	3.15E−10	2.4	22.7
GO:0006968~cellular defense response	9	2.72E−04	5.3	19.0
GO:0042060~wound healing	12	1.57E−04	4.1	15.7
GO:0065007~biological regulation	176	1.00E−10	1.5	14.9

naturally raises an important question: with which CRM target proteins do these partner proteins interact, or stated another way, which CRM targets *invoked* these partners? This would be of considerable interest in view of the hypothesis that perturbing the PPIs of these targets could possibly impact intracellular signaling by their partners and thereby have important toxicological consequences.

To find the targets that invoked the partners comprising the signaling cascades in Table 5, we returned to our original spreadsheet listing the interacting partners of each of the CRM target (Supplementary table S1) and re-sorted it by partner to reveal the target protein(s) associated with each of them (Supplementary table S2). When this was done for the individual KEGG pathways, the corresponding lists of target proteins were observed

to display considerable redundancy between them (Table 6A). We suggest that this redundancy gives an important clue to the potential "reach" that the covalent modification of a given target protein may have in terms of its potential to disrupt multiple PPIs and perhaps initiate toxicological outcomes for a cell. When this redundancy is compressed, simply by counting the number of times a given target either appears or is invoked by its partners (Table 6B), it is found that just seven target proteins are responsible for invoking 73 of the proteins listed in Table 5. These seven proteins are GSTP1, Hsp90AB1, HspA5, HspA8, HspD1, PEBP1 (also known as RKIP), and thioredoxin. This information will be combined with comparable information gleaned from similar analyses of the highest ranking GO categories as described below.

GO-BP Analysis of Targets and Partners

Because of the detailed and nested hierarchical structure of the GO classification system, the number of categories and subcategories is much larger than the number of KEGG pathways, but the approach to evaluating the sorting results is essentially the same. In the GO-BP category, 466 subcategory bins were flagged using low stringency criteria (Supplementary table S4). Ranking the bins by their CI and further culling for relevance to cytotoxicity reduced the number to 46 bins that we arranged into six groups (Table 7). As with the KEGG analysis above, each of these subcategory bins contained a number of proteins represented by a listing of their gene names. The total number of genes across all 46 subcategories was 1,670, but there was a great deal of redundancy among them, due again to the nested hierarchical nature of the GO classification scheme. Three of the largest subgroups in Table 7, namely groups 1, 2, and 6, were selected for further detailed analysis. Group 5 (kinases and signaling pathways), although large, was not analyzed further because of substantial redundancy with the KEGG analysis above. Group 3 (NO and inflammation) and group 4 (oxidative stress), while relevant for cytotoxicity, were not analyzed further because they were very small compared with the other groups.

In Table 7, group 1 consists of 15 subcategory bins containing a total of 522 genes of target or partner proteins. For convenience in analysis, and in parallel with the KEGG analysis above, these 15 subcategories were recombined into four subgroups according to the similarity (redundancy) of the genes included (data not shown). In the first subgroup, the total number of nonredundant genes was 68, among which only seven correspond to known target proteins, whereas the remaining 61 genes represent partner proteins. Analysis of these partners revealed that just 16 target proteins invoked all 61 partner proteins. Furthermore, as in the KEGG analysis above, these 16 target proteins are not *members* of these GO subcategories, but rather they interact with some of the members and thus flag the categories through the interacting partners. Analysis of the other three group 1 subgroups produced similar patterns of results regarding numbers of target versus partner proteins and redundancies among them. Pooling all the results for group 1 of Table 5 showed that overall only nine target proteins emerged from all four subgroups while nine more emerged from at least three of the four subgroups. In a similar fashion, analysis of group 2 from Table 5 (five subcategories listing 129 genes) revealed that seven target proteins occurred three or more times, whereas eight proteins occurred two or more times. Likewise for group 6 of Table 5 (12 subcategories listing 650 genes), 13 target proteins emerged from all five subgroups, whereas four more emerged from at least four of the five subgroups. This information was eventually combined with comparable information from the KEGG analysis (above) and the GO-MF analysis as described below.

GO-MF Category Analysis of Targets and Partners

In a parallel fashion, sorting our set of target plus partner proteins into the GO-MF classification initially yielded 119 subcategories (Supplementary table S5), but this number was reduced

TABLE 8
Top-Ranking GO-MF Subcategories (n = 18) and Their Statistical Parameters

GO-MF category	Count	p value	Fold enrichment	CI
GO:0008384-IkappaB kinase activity	3	1.28E–03	47.8	138
GO:0030911-TPR domain binding	3	1.28E–03	47.8	138
GO:0003756-protein disulfide isomerase activity	5	3.54E–05	23.9	106
GO:0030188-chaperone regulator activity	4	2.96E–04	27.3	96
GO:0030235-nitric-oxide synthase regulator activity	3	2.53E–03	35.8	93
GO:0030192-Hsp70/Hsc70 protein regulator activity	3	2.53E–03	35.8	93
GO:0051082-unfolded protein binding	18	2.82E–09	6.4	55
GO:0008092-cytoskeletal protein binding	36	7.34E–12	4.0	44
GO:0004879-ligand-dependent nuclear receptor activity	11	1.57E–06	7.6	44
GO:0003779-actin binding	28	2.19E–10	4.4	43
GO:0004709-MAP kinase kinase kinase activity	4	5.82E–03	10.6	24
GO:0019904-protein domain–specific binding	10	4.94E–04	4.3	14
GO:0031072-heat shock protein binding	7	2.14E–03	5.2	14
GO:0051087-chaperone binding	3	3.81E–02	9.6	14
GO:0005200-structural constituent of cytoskeleton	8	2.27E–03	4.4	12
GO:0004860-protein kinase inhibitor activity	4	3.10E–02	5.8	9
GO:0042287-MHC protein binding	3	8.24E–02	6.2	7
GO:0016209-antioxidant activity	5	2.93E–02	4.3	7

TABLE 9
Target Proteins With the Strongest Direct Links to Toxicity

Target protein name	Gene name	GO-MF	GO-BP	KEGG	Sum	Overall rank	Hit rank[a]
Actin, cytoplasmic 2; gamma-actin	ACTG1	41	6	0	47	1	12
Heat shock 70 kDa protein 8, Hsc 70	HSPA8	17	8	4	29	2	15
HSP90B, HSP 84	HSP90AB1	11	8	4	23	3	—
BIP, GRP 78, HSPA5	HSPA5	8	8	3	19	4	—
Tropomyosin 1 alpha chain	TPM1	10	5	0	15	5	6
Heat shock 60 KD protein 1 (chaperonin), HSP-60	HSPD1	3	8	3	14	6	3
Phosphatidylethanolamine-binding protein 1, Raf kinase inhibitor protein	PEBP1	4	6	3	13	7	—
Thioredoxin	TXN	2	8	3	13	8	—
Calreticulin	CALR	6	6	0	12	9	24
Albumin	ALB	4	7	0	11	10	4
GRP 94, tumor rejection antigen 1	HSP90B1	1	8	1	10	11	—
Dipeptidyl peptidase IV	DPP4	2	5	2	9	12	—
Glutathione S-transferase Pi	GSTP1	0	5	3	8	13	—
Protein disulfide isomerase A3, PDI A3, ER-60	PDIA3	2	5	1	8	14	2
Aldo keto reductase family 1, member C3	AKR1C3	0	6	0	6	15	—
Aldehyde dehydrogenase, mitochondrial	ALDH2	2	4	0	6	16	—
Endoplasmic reticulum protein 29	ERP29	2	4	0	6	17	—
Peroxiredoxin 6	PRDX6	3	3	0	6	18	16
Protease inhibitor 1, Alpha-1-antitrypsin 1-1	SERPINA1	2	4	0	6	19	22
ATP synthase, alpha	ATP5A1	0	5	0	5	20	23
Protein disulfide isomerase A1, PDI, Erp59	P4HB	4	1	0	5	21	1
ATP synthase, beta	ATP5B	0	4	0	4	22	18
Glutathione S-transferase Mu-1	GSTM1	0	3	1	4	23	—
Peroxiredoxin 2	PRDX2	0	3	1	4	24	—
Transthyretin	TTR	1	2	1	4	25	—
Aldolase B, fructose-bisphosphate	ALDOB	0	3	0	3	26	—
Fatty acid–binding protein 1	FABP1	2	1	0	3	27	27
Phosphoglycerate kinase 1	PGK1	0	3	0	3	28	—

Note. The ranking is based on taking PPIs into account as described in the text and in Figure 1. The overall rank is based on the sum of the KEGG, GO-BP, and GO-MF rankings (Tables 4, 7 and 8, respectively). The ranking by number of hits given in Table 2 is repeated here for comparison.
[a]Hit rank scores are from Table 2.

to 18 (Table 8) by considering CI values and relevance to cytotoxicity. These 18 subcategories contained a total of 158 genes representing target and partner proteins, but the 158 genes corresponded to only 44 nonredundant proteins, among which only 5 were actual CRM target proteins, whereas 39 corresponded to partner proteins. Analysis of these partner protein genes as a single group yielded a list of just 16 targets that were responsible for invoking all 39 partner proteins among the 18 subcategories. Of the 21 target proteins identified by this search, all had already been identified as relevant and potentially important through the KEGG or GO-BP searches described above. Thus, although the GO-MF analysis revealed no new proteins of interest, it clearly reinforced the likely importance of those already identified using the KEGG and GO-BP approaches.

PPIS AS TARGETS FOR CRMS

Examples of the modulation of PPIs by small changes at a single residue is a well-known phenomenon. For example, mutation of the single cysteine residue in murine Hsp25 to alanine interferes with its protective functions by preventing its dimerization (Diaz-Latoud et al., 2005). Covalent modification of Cys49 in the cytosolic domain of microsomal GST-1 by electrophiles activates this protective enzyme and thereby "acts as an antenna for the detection of chemical or oxidative stress" (Busenlehner et al., 2004). In nonstressed cells, the activity of JNK is kept low by its 1:1 complexation to GST-pi (which has four reactive sulfhydryl groups near its active site), but this suppression is relieved by oxidative stress, by the addition of GSH analogs that bind to GST-pi, and by ethacrynic acid (a Michael acceptor that S-alkylates GST-pi) (Townsend and Tew, 2003). These and other examples suggest that covalent modification of proteins by CRMs could affect cellular homeostasis by modulating or interfering with PPIs.

Of course, not all CRM-induced changes in PPIs are necessarily deleterious. The Keap1-Nrf2-ARE pathway provides a clear example (Dinkova-Kostova et al., 2005; Goldring et al., 2004; Kwak et al., 2004; Nguyen et al., 2003). This pathway regulates several protective mechanisms including the upregulation of conjugating and antioxidant enzymes, anti-inflammatory responses, the molecular chaperone/stress-response system, and the ubiquitin/proteasome system. The protein Keap1 has 27 cysteine sulfhydryl groups, many of

which are reactive toward electrophiles (Hong et al., 2005). Modification of sulfhydryl groups on Keap1 causes the transcription factor Nrf2 to dissociate, allowing it to migrate to the nucleus and trigger the expression of genes favorable to cell defense that are under the control of the antioxidant response element (ARE, also called the electrophile response element). Although the CRM-induced changes in the activity of the Keap1-Nrf2-ARE pathway appear to be protective, alteration of other PPIs could, in principle, just as easily lead to detrimental effects on the cell.

SUMMARY AND CONCLUSIONS

The names and frequencies of occurrence of the target proteins emerging from the GO-MF, GO-BP, and KEGG analyses are compiled in Table 9, and a diagram of the interactions of 12 "well-connected" targets with partners is given in Figure 2. It should be noted that the individual frequencies of occurrence of the target proteins listed in Table 9 could have been made to appear higher, and the resulting rank ordering could perhaps have changed slightly if we had included even more subsets of GO and KEGG pathways in our analysis. However, the effect of "stacking the deck" like this would have been to increase redundancy among the bins and decrease stringency from a statistical and a biological relevance perspective. In our view, doing so would have added no further value or precision to the result, which is a small subset of CRM target proteins, winnowed from many, by an objective process that is also mindful of relevance to toxicity. This process explicitly took into account the hypothesis that disruption of PPIs (Fig. 2) by covalent binding contributes significantly to CRM-induced cytotoxicity. The implication is that target proteins near the top of Table 9 may therefore be more important than many other target proteins in generating biological responses to cytotoxic CRMs.

The relative rankings of proteins generated by the method described in this article (Table 9) are obviously considerably different from those given in Table 2, which considered only the number of times a given protein is reported as a target in the literature. It must be noted that the literature on target proteins is actually rather limited, and the data correspondingly sparse, for a number of reasons. For example, low abundance proteins and proteins with low levels of adduction are likely not to be reported because they are missed (i.e., not detected) in autoradiograms of 2D gels or because they fail to be detected as identifiable peptides in digests of excised spots (Koen et al., 2007). Another major reason that the literature on CRM targets is incomplete is the general lack of systematic searching for adducts of a given drug or chemical across multiple species, or tissues or cell lines. The vast majority of the 46 chemicals for which target proteins have been identified have been studied in only one species or one type of cell. Consequently, many of the proteins listed in the TPDB are reported only in a single study involving a single chemical and a single biological system. This can easily be seen by

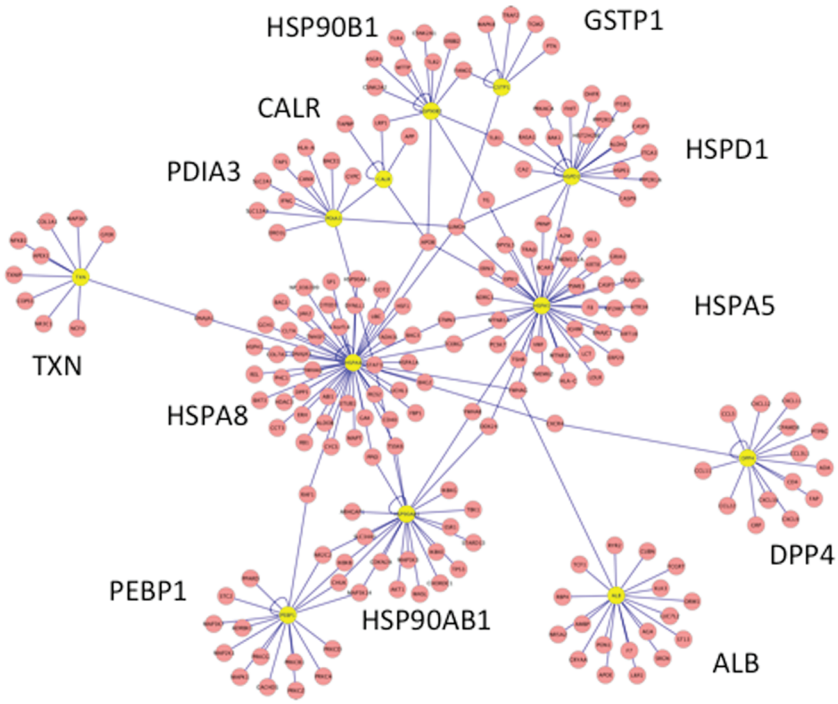

FIG. 2. PPIs among 12 key CRM target proteins and 194 interacting partner proteins most strongly linked to cytotoxicity. The 12 target proteins (yellow nodes) are among the top 14 proteins listed in Table 9. Only two target proteins interact with each other directly (PDIA3-CALR), whereas 14 partner proteins interact directly with two or more target proteins.

using the "commonality matrix" and "rank by hits" functions built into the TPDB (2013).

The comprehensive bioinformatic analysis described in this report is an attempt to transcend deficiencies in the CRM-target literature and differences in experimental design choices among the numerous published studies and to distill out a small number of proteins that may have the highest relevance to CRM-dependent cytotoxicity. The result is a rather different view of which target proteins might be most important for CRM-induced cytotoxicity compared with any previous analysis. For example, only 7 of the top 15 proteins from Table 2 are included among the top 15 proteins in Table 9, and 1 of those 7, calreticulin, was ranked quite low in Table 2. One possible reason that calreticulin is ranked relatively low in Table 2 is that being a relatively acidic protein (pI = 4.29) it is often missed in 2DGE analyses of target proteins. On the other hand, its relatively higher listing in Table 9 stems in part from the fact that it has at least 34 known partner proteins and many of them sort into KEGG pathways and GO categories that relate in a direct way to cytotoxicity. Thus, calreticulin can be considered to have a high potential "reach" into a number of biochemical pathways. Conversely, the two highest ranking proteins in Table 2, protein disulfide isomerase A1 and A3, are ranked somewhat lower in Table 9, probably because they have many fewer known partners that also happen to sort into KEGG pathways and GO categories of intrinsically lower relevance to acute cytotoxicity.

In conclusion, we suggest that taking partner proteins into account when examining collections of CRM target proteins is useful and may prove to be important in facilitating mechanistic understanding of reactive metabolite cytotoxicity. Our results also support the hypothesis that disruption of PPIs can be at least a major contributor if not a major mechanism of CRM-induced acute cytotoxicity. Additional experimental work will obviously be required to examine in detail the roles of individual proteins highlighted through this bioinformatic approach.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

FUNDING

U.S. National Institutes of Health (GM-21784 to R.P.H.).

ACKNOWLEDGMENTS

We thank Dr Jeffrey Staudinger, Dr John Karanicolas, and Dr Todd Williams for helpful discussions during the writing of this manuscript.

REFERENCES

Busenlehner, L. S., Codreanu, S. G., Holm, P. J., Bhakat, P., Hebert, H., Morgenstern, R., and Armstrong, R. N. (2004). Stress sensor triggers conformational response of the integral membrane protein microsomal glutathione transferase 1. *Biochemistry* **43,** 11145–11152.

Diaz-Latoud, C., Buache, E., Javouhey, E., and Arrigo, A. P. (2005). Substitution of the unique cysteine residue of murine Hsp25 interferes with the protective activity of this stress protein through inhibition of dimer formation. *Antioxid. Redox Signal*. **7,** 436–445.

Dinkova-Kostova, A. T., Holtzclaw, W. D., and Kensler, T. W. (2005). The role of Keap1 in cellular protective responses. *Chem. Res. Toxicol*. **18,** 1779–1791.

Druckova, A., Mernaugh, R. L., Ham, A. J., and Marnett, L. J. (2007). Identification of the protein targets of the reactive metabolite of teucrin A in vivo in the rat. *Chem. Res. Toxicol*. **20,** 1393–1408.

Evans, D. C., Watt, A. P., Nicoll-Griffith, D. A., and Baillie, T. A. (2004). Drug-protein adducts: An industry perspective on minimizing the potential for drug bioactivation in drug discovery and development. *Chem. Res. Toxicol*. **17,** 3–16.

Fang, J., Koen, Y. M., and Hanzlik, R. P. (2009). Bioinformatic analysis of xenobiotic reactive metabolite target proteins and their interacting partners. *BMC Chem. Biol*. **9,** 5.

Goldring, C. E., Kitteringham, N. R., Elsby, R., Randle, L. E., Clement, Y. N., Williams, D. P., McMahon, M., Hayes, J. D., Itoh, K., Yamamoto, M., *et al.* (2004). Activation of hepatic Nrf2 in vivo by acetaminophen in CD-1 mice. *Hepatology* **39,** 1267–1276.

Hanzlik, R. P., Fang, J., and Koen, Y. M. (2009). Filling and mining the reactive metabolite target protein database. *Chem. Biol. Interact*. **179,** 38–44.

Hong, F., Freeman, M. L., and Liebler, D. C. (2005). Identification of sensor cysteines in human Keap1 modified by the cancer chemopreventive agent sulforaphane. *Chem. Res. Toxicol*. **18,** 1917–1926.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res*. **37,** 1–13.

Ikehata, K., Duzhak, T. G., Galeva, N. A., Ji, T., Koen, Y. M., and Hanzlik, R. P. (2008). Protein targets of reactive metabolites of thiobenzamide in rat liver in vivo. *Chem. Res. Toxicol*. **21,** 1432–1442.

Koen, Y. M., Gogichaeva, N. V., Alterman, M. A., and Hanzlik, R. P. (2007). A proteomic analysis of bromobenzene reactive metabolite targets in rat liver cytosol in vivo. *Chem. Res. Toxicol*. **20,** 511–519.

Koen, Y. M., Williams, T. D., and Hanzlik, R. P. (2000). Identification of three protein targets for reactive metabolites of bromobenzene in rat liver cytosol. *Chem. Res. Toxicol*. **13,** 1326–1335.

Krishna, R. G. and Wold, F. (1998) Posstranslational modifications. In *Proteins: Analysis and design (Angeletti, R. H.,* Eds.) pp 121–207.

Kwak, M. K., Wakabayashi, N., and Kensler, T. W. (2004). Chemoprevention through the Keap1-Nrf2 signaling pathway by phase 2 enzyme inducers. *Mutat. Res*. **555,** 133–148.

Liebler, D. C. (2006). The poisons within: Application of toxicity mechanisms to fundamental disease processes. *Chem. Res. Toxicol*. **19,** 610–613.

Liebler, D. C., and Guengerich, F. P. (2005). Elucidating mechanisms of drug-induced toxicity. *Nat. Rev. Drug Discov*. **4,** 410–420.

Nguyen, T., Sherratt, P. J., and Pickett, C. B. (2003). Regulatory mechanisms controlling gene expression mediated by the antioxidant response element. *Annu. Rev. Pharmacol. Toxicol*. **43,** 233–260.

Park, B. K., Kitteringham, N. R., Maggs, J. L., Pirmohamed, M., and Williams, D. P. (2005). The role of metabolic activation in drug-induced hepatotoxicity. *Annu. Rev. Pharmacol. Toxicol*. **45,** 177–202.

Park, B. K., Laverty, H., Srivastava, A., Antoine, D. J., Naisbitt, D., and Williams, D. P. (2011). Drug bioactivation and protein adduct formation

in the pathogenesis of drug-induced toxicity. *Chem. Biol. Interact*. **192,** 30–36.

Reinders, J., and Sickmann, A. (2007). Modificomics: Posttranslational modifications beyond protein phosphorylation and glycosylation. *Biomol. Eng*. **24,** 169–177.

Snyder, R. (2011). Origins of the International Symposium on Biological Reactive Intermediates and some thoughts on homeostasis and the biological significance of widespread covalent binding of biological reactive metabolites. *Chem. Biol. Interact*. **192,** 3–7.

Townsend, D. M., and Tew, K. D. (2003). The role of glutathione-S-transferase in anti-cancer drug resistance. *Oncogene* **22,** 7369–7375.

TPDB (2013). Reactive Metabolite Target Protein Database. Available at: http://tpdb.medchem.ku.edu:8080/protein_database/. Accessed February 2, 2013.