View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

Accepted Manuscript

Method for suppressing non-specific protein interactions observed with affinity resins

J.S. Rees, K.S. Lilley

PII: S1046-2023(11)00120-4
DOI: 10.1016/j.ymeth.2011.06.004
Reference: YMETH 2749

To appear in: *Methods*

Received Date: 22 December 2010
Revised Date: 2 June 2011
Accepted Date: 6 June 2011

Please cite this article as: J.S. Rees, K.S. Lilley, Method for suppressing non-specific protein interactions observed with affinity resins, *Methods* (2011), doi: 10.1016/j.ymeth.2011.06.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

1 Method for suppressing non-specific protein interactions observed with affinity resins.

2

3 J.S. Rees & K.S. Lilley*

4

5 Cambridge Centre for Proteomics, Department of Biochemistry, University of
6 Cambridge, Tennis Court Road, Cambridge, CB2 1QR, UK.
7 jsr31@cam.ac.uk and ksl23@cam.ac.uk (*corresponding author)

8

9 **Abstract**
10 Previous high throughput data analysis from several different approaches to affinity
11 purification of protein complexes have revealed catalogues of contaminating proteins
12 that persistently co-purify. Some of these contaminating proteins appear to be
13 specific to one particular affinity matrix used or even to the artificial affinity tags
14 introduced into endogenous proteins for the purposed of purification.
15 A recent approach to minimising non-specific protein interactions in high throughput
16 screens utilises pre-equilibration of affinity surfaces with thiocyanate anions to reduce
17 non-specific binding of proteins. This approach not only reduces the effect of
18 contaminating proteins but also promotes the enrichment of the specific binding
19 partners. Here, we have taken this method and adapted it in an attempt to reduce the
20 abundance of common contaminants in affinity purification experiments. We found
21 the effect varied depending on the bait used, most likely due to its endogenous
22 abundance.

23

24 **Keywords.** Affinity purification, non-specific, contaminants, thiocyanate, Mass
25 Spectrometry

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

26 **1. Introduction**

27

28 The characterisation of native protein interactions is essential for our understanding of
29 the processes which underlie biological functions. In order to gain a comprehensive
30 knowledge of multi-component protein complexes it has been necessary to develop
31 and utilise high throughput methods which allow the identification of genuine
32 interaction partners. This has lead to the inception of the field of interactomics, a
33 rapidly growing field with numerous different approaches developed to allow the
34 characterisation of proteins within functional complexes. Many of these approaches
35 involve the affinity capture of a bait protein, either by its interaction with a specific
36 antibody or by interaction of an engineered component such as a short protein epitope
37 tag or full length fusion protein. After affinity capture, identification of the bait and its
38 interacting partners are generally achieved by mass spectrometry. Approaches such as
39 the tandem affinity purification method (TAP) allow high-throughput screening of
40 interactomes in multicellular organisms [1]. Here, bait proteins are tagged with two
41 affinity tags and purification of the tagged bait and its interacting partners is then
42 carried out using the affinity properties of each tag sequentially. Another recent
43 technique, iPAC (interactomes by parallel affinity capture), favours parallel
44 purifications of a multiply tagged protein to increase yields of purified complexes as
45 tandem approaches often result in very low recoveries of protein complex components
46 after multiple sequential application and elutions from affinity matrices [2]. In all
47 these approaches, conditions are utilised to minimise the sampling of contaminants
48 such as stringent washing of affinity matrices before specific elution of the bait and its
49 binding partners and occasionally the implementation of exclusion lists of ions
50 associated with common contaminating proteins during mass spectrometric analysis.
51 Despite these precautions, contaminants that have high affinity to single or multiple
52 resins continue to be a problem in blocking available binding sites for the tagged
53 protein(s) thus resulting in low recovery yields of genuine interacting partners.
54 Moreover, these proteins can dominate mass spectrometric analyses, usually in the
55 form of peptides generated upon proteolytic digestion of eluted complex components
56 prior to analysis. Without appropriate experimental designs it can be challenging to
57 differentiate between genuine interacting partners and contaminants. One method
58 which aids differentiation involves the use of quantitative approaches where a
59 negative control such as a system without a tagged bait is applied to the same affinity

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

60 matrix as the tagged version and the abundance of eluted proteins compared and the
61 enrichment of genuine partners is established [3,4]. Trinkle-Mulchay and co-workers
62 used this approach to quantify proteins that non-specifically bound to a GFP-Trap
63 resin used for isolating and purifying GFP-tagged proteins and their binding partners
64 from complex mixtures [5,6].
65
66 A complementary approach is to minimise binding of contaminants prior to purifying
67 proteins of interest. High throughput chip arrays often use a blocking system that
68 enable genuine binders, that have higher affinity, to preferentially bind to specific
69 binding sites. This approach is especially useful for isolating proteins of low
70 abundance from complex mixtures from whole cell lysates. Recently, Richens and co-
71 workers used thiocyanate, a member of the Hofmeister series [7,8] in order to reduce
72 the non specific binding of abundant proteins such as albumin on label free protein
73 arrays. Thiocyanate is a relatively large anion which has a very high entropy of
74 hydration. It is thought to disrupt non-specific interactions of proteins by modulating
75 the structure of water in surrounding interacting regions of macromolecules and thus
76 disrupts selectively the non-polar effects that facilitate non-specific binding events
77 [8,9]. Richens and colleagues demonstrated enrichment of low abundance proteins,
78 often 10 orders of magnitude lower than that of albumin, the predominant component
79 of serum, when binding assays were carried out in the presence of thiocyanate anions
80 [9].
81 Here we apply the approach of pre-treating affinity matrices designed for the
82 purification of tagged protein baits and their interacting partners with thiocyanate
83 containing buffers and demonstrate a reduction in the co-purification of some of the
84 common contaminants regularly described in the literature. In taking this approach
85 we facilitate the maintenance of transient or short lived interactions. We demonstrate
86 that thiocyanate pre-treatment of affinity binding matrices and, more importantly,
87 inclusion in the binding step is efficient at increasing bait peptide identification as
88 well as reducing non-specific binding events within the iPAC protocol.
89 The objective of the study presented here is to assess the effect of inclusion of
90 thiocyanate ions in affinity purifications using multiple affinity resins and to reduce
91 the numbers of non-specific contaminants allowing surveying of lower abundance
92 proteins in complex mixtures.
93

ACCEPTED MANUSCRIPT

2. Methods

2.1 Tagged lines

A number of affinity tagged *Drosophila melanogaster* lines from the FlyProt collection (Kyoto stock centre, www.flyprot.org) were randomly selected for affinity purifications to analyse both genuine and non-specific binding proteins. These comprised a tandem triple tag of FLAG-Strep-YFP-Strep with the former used for affinity purifications and the YFP for visual assessment (Figure 1A). We also used non-tagged w118 control flies to determine non-specific binding proteins to identify proteins that bind to the resin material and a positive control Vha55-YFP-Strep to test the effect of the treatment.

2.2 Assessing suitability of thiocyantate anions

All affinity purifications of triple tagged proteins from *D. melanogaster* embryo lysates were performed as described in Rees *et al* [2] with the following additions and modifications. To determine the optimal buffering conditions and chaotropic anion concentration a pilot study was performed with a well characterised, high abundance, bait and a non-tagged control. Lysates were prepared in Veraksa buffer [10] including protease inhibitor cocktail (Roche) with or without 50-500 mM NaSCN (thiocyanate) and 2.85 mM PBS and incubated with either ANTI-FLAG M2 MAb (Sigma) or *Strep*-Tactin (IBA) sepharose resins that were pre-equilibrated in 50-500 mM thiocyanate and 2.85 mM PBS. After binding for one hour resins were washed three times for 10 min in Veraksa buffer to remove non specific binders. *Bona fida* bound native protein complexes were eluted twice each with either 50 l (100 g/ml) FLAG peptide (Sigma) or 50 l of 10 mM Biotin (Sigma) respectively in Veraksa buffer containing protease inhibitor cocktail for 30 min at 4ºC on a rotary mixer.

.Pooled eluates and non-binding fractions were firstly analysed by immunoblot to detect recovery of the bait and actin abundance. The workflow is summarised in Figure 1B.

Suggested location for Figure 1 here.

2.3 Protein identification by mass spectrometry.

Total eluates were partially resolved by SDS-PAGE, stained with Coomassie, excised, reduced in 2 mM DTT for 1 hour at RT and alkylated in 10 mM iodoacetamide for 30

ACCEPTED MANUSCRIPT

127 min at RT. Proteins were digested with 2 g sequencing grade trypsin (Promega) for 1
128 hour at 37ºC, then a further 2 g for overnight digestion to maximise complete
129 digestion of complex mixtures.
130 Digests were prepared in 0.1% Formic acid and analysed in a single run by Mass
131 Spectrometry (MS) to identify all proteins eluted. MS was performed as described in
132 Rees *et al* [2] but without actin containing exclusion lists, in order to measure the
133 abundance of actin with and without treatment. The Orbitrap was operated in data
134 dependent mode, acquiring an MS scan and two subsequent MS/MS measurements
135 with a precursor dynamic exclusion of 0.3 Da.
136 Peak lists were generated using Bioworks Browser version 3.3.1 (2007). Resulting
137 fragment masses (MS/MS) were searched using the Mascot version 2.2 (Matrix
138 Science) search engine against an in house database comprising the FlyBase *D.*
139 *melanogaster* genome (version 5.9) totalling 21064 proteins, plus the FASTA
140 sequence for YFP as a secondary confirmation of the presence of the tagged protein.
141 Parameters included a precursor mass tolerance of 1.0 Da and fragment ion mass
142 tolerance of 0.8 Da, 2 missed cleavages and methionine oxidation as variable and
143 carbamidomethylated cysteine as fixed modifications. The decoy database option,
144 comprising a scrambled *D. melanogaster* database *in silico* digested that generates a
145 similar number of the same sized peptides, was selected to automatically calculate the
146 protein false discovery rate (FDR). Stringent parameters were used to ensure
147 accuracy in the datasets. For example, proteins with single peptide hits were
148 eliminated. MS samples were run once or twice if the bait was of particularly low
149 abundance.
150 Resulting proteins lists were exported and compared using the ProteinCenter software
151 (Thermo).
152
153 *2.4 Interaction validation.*
154 To determine if the protein interaction partners we observed are genuine, we used
155 FlyMine search queries (www.flymine.org) to mine the public interaction databases,
156 such as IntAct. Binary search queries within *Drosophila melanogaster* interaction
157 datasets identify proteins within our list that have had reported interactions and binary
158 search queries in orthologous interaction datasets allow us to potentially highlight

ACCEPTED MANUSCRIPT

159 observed interactions in orthologous species, such as yeast, worm, human, that have
160 not yet been detected in flies.
161
162 **3. Results**
163 *3.1 Parallel affinity purifications.*
164 Previous purifications of triple tagged *Drosophila melanogaster* proteins have used
165 parallel methods where soluble extracts are split and half purified using Strep-Tactin
166 resins and the other half FLAG monoclonal antibody (MAb) resins, all eluates tryptic
167 digested and analysed by Mass Spectrometry (MS) and the resulting protein lists
168 compared *in silico.* Almost 250 proteins' interacting partners have been characterised
169 by this method and in almost all, actin was a non-specific interactor [2].
170
171 *3.2 Pre-treatment of affinity resins.*
172 Equilibration of affinity resins is a pre-requisite for efficient binding of target
173 proteins. Pre-binding Strep-Tactin and FLAG MAb resins with thiocyanate should
174 enable selective competition of binding of genuine tagged proteins so we tried pre-
175 equilibrating, post washing, with concentrations of 50-500 mM NaSCN. MS analysis
176 of the resulting eluates showed little difference in the non-specific binding to the
177 resins according to protein lists generated from MS analysis (data not shown).
178
179 *3.3 Thiocyanate ions improve the specific binding of proteins to affinity resins.*
180 The next approach was to include sodium thiocyanate in the binding mixture. To
181 determine if the effect of the addition of thiocyanate ions in affinity purification of
182 protein complexes is beneficial, we first used non-tagged control fly embryo lysates to
183 identify all proteins that bind non-specifically to the FLAG affinity resins. Several
184 concentrations of the sodium thiocyanate were utilised ranging 50-500 mM. The most
185 efficient concentration of the thiocyanate anions utilised seemed to be relatively broad
186 as over a wide concentration tested, all gave the same protein identification lists
187 therefore 100mM thiocyanate was used in further experimentation.
188 This proof of principle experiment that used control lines, where proteins should not
189 bind resin, demonstrated that the addition of 100 mM thiocyanate reduces the number
190 of non-specific proteins eluted from FLAG resins by a specific FLAG peptide
191 compared to purifications without thiocyanate, from 37 to 32 in one test and from 30
192 to 24 in a biological replicate (Supplementary Table 1).

ACCEPTED MANUSCRIPT

193 We also used a well characterised protein, Vha55, a subunit of the mitochondrial
194 VATPase complex to demonstrate enrichment of known and predicted binding
195 partners in the presence of thiocyanate. Of the 175 and 85 *D. melanogaster* proteins
196 detected with or without thiocyanate respectively, 73 and only 24 (42% and 28%)
197 were remaining after removal of the corresponding W- negative controls. Importantly
198 FlyMine search queries revealed 22 proteins were identified as known or predicted
199 interacting partners for thiocyantate treated compared to only 2 for the non treated
200 sample (Supplementary Figure 1) demonstrating that thiocyanate enriches for genuine
201 interacting partners. A further 35 proteins from thiocyanate-treated eluates were
202 known or predicted but were also found in negative control samples with the majority
203 enriched for in the positive compared to negative based on peptide numbers, sequence
204 coverage and spectral counts (Supplementary Table 2).
205 Based on these observations we then sampled 8 different triple- FLAG-Strep-YFP-
206 Strep-tagged protein lysates to determine if the inclusion of thiocyanate would
207 improve the binding and numbers of specific binders and reduce the number of non-
208 specific binders, both to the resin and the bait. Immunoblots showed, in many cases,
209 an increase in the yield of bait, as detected by anti GFP antibody (Supplementary
210 Figure 2A).
211
212 Mass spectrometry analysis showed increased numbers of bait peptides identified in
213 the presence of thiocyanate for 9 of the 10 tagged proteins tested with a range of a 16-
214 >100% increase, the average being 43.5% (Table 1). This was similar for the YFP
215 peptides also generated from the bait protein. This trend was also observed in the %
216 sequence coverage of the bait protein. Whilst proteins purified using Strep resin had
217 higher numbers of peptides, the effect of thiocyanate was more dramatic for FLAG
218 purified proteins and in general the addition of thiocyanate was beneficial for
219 increasing the binding of bait to both FLAG and Strep resins. Mascot or emPAI scores
220 were more ambiguous with respect to the effect of thiocyanate ions.
221 **Suggested location for Table 1 here.**
222
223 The protein lists generated by MASCOT were compared *in silico* to analyse proteins
224 eluted in the presence or absence of thiocyanate and Venn diagrams were used to
225 show the overlap after removal of negative control proteins (Figure 2, upper panel)
226 and 'non-specific' binders (identified in >20% of all interaction lists irrespective of

ACCEPTED MANUSCRIPT

227 the bait, as shown in Figure 3A and Supplementary Table 3A&B) (Figure 2, lower
228 panel). In all cases there were different proteins identified by each unique experiment.
229 **Suggested location for Figure 2 here.**
230
231 Protein lists for thiocyanate treated lysates had reduced abundance of actin, and fewer
232 proteins were categorised as non-specific, particularly in FLAG pulldowns.. The
233 remaining proteins contained uncharacterised proteins and a proportion of proteins
234 similar to those found to be non-specific binders such as heat shock proteins and
235 tubulins that were not frequent enough to be included in the 'non-specific' lists. After
236 identifying the recurring non-specific members, lists from non-treated lysates also
237 contained a high proportion of other ribosomal proteins, metabolic proteins and
238 uncharacterised proteins that are unlikely to be genuine interactors of the bait.
239 Annotated interaction lists are shown in Supplementary Table 4A.
240 When analysing the interacting proteins, very few of the proteins studied had any
241 published interaction data in *Drosophila* (Supplementary Table 4B). Using FlyMine
242 we did observe novel interactors in eluates both from inclusion and exclusion of
243 thiocyanate. For example, the bait Pop2 (CPTI 2818) interacts with proteins Not1 and
244 twin in yeast. These were both found in FLAG pulldowns but in the less effective
245 Strep pulldown, were only found with the inclusion of thiocyanate (Supplementary
246 Table 4A). In addition, three known contaminants were found only in untreated
247 samples. Comparing our datasets with public datasets using FlyMine, many of our
248 interactors were seen in affinity purification studies in other species and some did
249 indeed complement Y2H studies (Supplementary Table 4B). It appears that
250 thiocyanate is useful in recovering some binding partners *in vivo* in some baits.
251
252 For analysing non-specific binders in more detail all of the protein lists from the 8
253 different baits and controls, comprising 25 different experiments, were combined *in*
254 *silico* using ProteinCenter software to determine frequently occurring proteins. The
255 most frequently occurring proteins were indeed identified in previous studies and are
256 illustrated in Figure 3A and Supplementary Tables 3 A&B. Several proteins were
257 eliminated in the presence of thiocyanate; CG9436, Tm1 (2 isoforms), RpL23A, Tm2,
258 and Ald. A further 7 proteins had between 30-50% reductions in occurrence of
259 appearances.
260 **Suggested location for Figure 3 here.**

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

261

262 A semi-quantitative analysis was performed to see the overall effect of the presence of
263 thiocyanate by plotting the average number of peptides observed in the 20 most
264 common contaminants from negative controls and tagged proteins both in the
265 presence and absence of thiocyanate ions, determined in Figure 3A, by globally
266 analysing all peptides from these 20 proteins generated in all 25 experiments (Figure
267 3B). The more commonly occurring proteins and peptides were also in negative
268 control samples suggesting that these are resin specific contaminants irrespective of
269 the bait protein used. Fewer peptides were observed for some structural scaffold
270 proteins such as actins and tubulins. From the top 20 contaminants' list 8 high
271 abundance proteins generated fewer peptides on average, ranging from a 33-2%
272 decrease, although 9 proteins had increased peptides whilst three showed no change
273 (Figure 3B). In addition, of the 72 proteins seen more than five times (≥20%
274 frequency) in the 25 pooled experiments, 35 proteins had reduced numbers of
275 observed peptides eluted from FLAG resins whilst only 29 had increased peptides in
276 the presence of thiocyanate (n=16) (Supplementary Tables 3A&B). For Strep resins
277 27 proteins had increased peptides compared to 31 that had lower numbers in the
278 presence of thiocyanate (n=9). We also looked at the changes in the average
279 percentage sequence coverage for all proteins occurring more than once in the 25
280 experiments. The heat map (Supplementary Table 3B) clearly shows that the changes
281 in protein sequence coverage mimic the changes in peptides and protein frequencies,
282 confirming that analysing changes in peptide numbers is a good tool for assessing the
283 effectiveness of thiocyanate. Mascot scores were not as reliable, as can be seen from
284 the heat map of the bait proteins (Supplementary Table 3B) as these varied widely
285 amongst biological and technical replicates.
286 Therefore we have not reduced all contaminant proteins but nevertheless, we have
287 improved the coverage of the bait and identified novel interacting partners for some
288 baits that are not present in controls or known contaminant lists.
289

290 **4. Discussion.**

291 A potential improvement to affinity purifications was performed and analysed with a
292 view of minimalising non-specific interactions pre-MS analysis. Thiocyanate ions
293 have been known to reduce binding of non-specific plasma proteins to protein chips

ACCEPTED MANUSCRIPT

294 and we found that the presence of thiocyanate in the binding step in our experiments
295 enriched for known binding partners in our positive control but did not necessarily
296 reduce the level of known contaminants such as actin and yolk proteins, although
297 these contaminants did not dominate in this sample. When testing poorly
298 characterised and/or lower abundance baits, in most examples we observed
299 enrichment of the bait and in one example we saw enrichment for known binding
300 partners unique to thiocyanate treatment. However many novel proteins identified
301 were unique to thiocyanate treated samples and will need further testing. In terms of
302 reducing non specific binding contaminants in some samples we observed varying
303 levels of reduction of one of the most abundant proteins, actin, but it varied depending
304 on the bait.
305 We tried the recommended concentration discussed by Richens and found it to be
306 effective in some of our experiments for minimalising non-specific binding of
307 structural proteins. However, by reducing these, we did observe increased binding of
308 other non-specific proteins such as the yolk proteins, probably because of the
309 increased sampling of lower abundance contaminants, but the reduction of actin and
310 other scaffold proteins outweighed the marginal increase in other contaminants. We
311 think this is a reasonable trade-off especially as we are aware of these common
312 contaminants in previous studies [2].
313 The effect of thiocyanate is clearly bait abundance specific and may be more
314 pronounced with lower abundance baits if MS data dependant exclusion lists for other
315 non-specific binders were used in parallel and moreover, our method has assisted in
316 identifying proteins which are the most desirable to exclude such as the yolk proteins
317 [2]. It is important in mass spectrometry experiments to be very careful when
318 excluding masses for analysis, as a too large an exclusion list will result in under
319 sampling of the proteome. Utilisation of this thiocyanate method to show which
320 proteins persistently co-fractionate, assists in choosing the most effective exclusion
321 parameters.
322 Future work would include investigating other members of Hofmeister series around
323 position of thiocyanate and see how they fare.
324
325
326 **Figure legends.**
327

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

328 **Figure 1. A.** Construct used to generate tagged endogenous proteins. Pl= Plasmid
329 region, PB *PiggyBac* Inverted repeats, P= P-ends, SA= Splice Acceptor, FLAG=
330 FLAG tag, S= *Strep*II tags, SD= Splice Donor. **B.** Workflow for the purification of
331 *Drosophila melanogaster* tagged proteins with and without the inclusion of
332 thiocyanate at various steps indicated. S= Strep tag, w= w118 control line, x and y are
333 tagged test lines. R= resin and E= eluate containing purified protein.
334
335 **Figure 2.** The effect of thiocyanate in reducing non specific protein binding to FLAG
336 and Strep affinity resins. **A-C.** Venn diagrams to show control proteins (Wf+&-) and
337 eluted proteins from pulldowns in the presence (+) or absence (-) of thiocyanate.
338 Subset Venn in **A'** is the merged negative control data. **D-F.** Venn diagrams showing
339 each bait with its respective negative control data subtracted and the proportion of the
340 data that is present more than 20% in all samples analysed 'non-specific' (list of
341 proteins in Supplementary Table 3A).
342
343 **Figure 3. A.** The numbers, and some identifications of proteins identified in 25
344 experiments using 8 protein baits. Single hit proteins have been excluded. Shaded
345 blocks indicate proteins occurring in >20%-100% frequently occurring interaction
346 lists from all pulldown experiments (including negative controls) thus defined as
347 'contaminants'. These proteins are listed and detailed in Supplementary Table 3A.
348 Dashed line shows the threshold we define our cut-off for contaminants. Proteins in
349 bold were not frequently occurring in negative control samples so are likely to be bait
350 specific and not resin specific (from Supplementary Table 1). **B.** An analysis of the
351 numbers of peptides generated from the 20 most abundant proteins (and their
352 frequencies in all interaction lists) observed in Figure 3A in the presence (grey bars)
353 or absence (black bars) of thiocyanate. The most notable differences are in the
354 scaffold proteins, actins and tubulins. The % average peptide count decreases,
355 compared with no treatment, are displayed above the bars.
356
357 **Table 1.** Mascot data from Mass Spectrometry analysis of 8 tagged (bait) proteins
358 purified in the presence (+) or absence (-) of thiocyanate. All FDRs are below 5%.
359 C=cytoplasmic, m=membrane, n=nuclear and unk=unknown. PG/PC are distinct
360 protein isoforms G and C respectively.
361

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

362 **Supplementary Figure 1.** Venn diagram to show proteins identified in the Vha55
363 positive and W118 negative controls with (+) and without (-) thiocyanate treatment.
364 Subsets highlight the number of known and predicted Vha55 interactors.
365
366 **Supplementary Figure 2. A.** Western blot to show the yields of bait recovered from
367 pulldowns in the presence (+) or absence (-) of Thiocyanate. Wf = w118 control line
368 pulled down with Flag (f), Heph (CPTI2584) and Cat (CPTI2786) are two FLAG-
369 Strep-GFP tagged proteins with their CPTI identifiers. s= soluble extract, e= eluate.
370 Black arrows indicate enrichment of bait. **B.** The number of peptides identified from a
371 single Mass Spectrometry analysis of the bait and different actin proteins (range of
372 peptides) in eluates from the inclusion (+) or absence (-) of thiocyanate.
373
374 **Supplementary Table 1.** Mass Spectrometry data from non-tagged controls to
375 determine proteins being eluted non-specifically from FLAG and Strep resins and
376 analysis of the peptides generated. Wf1+; W=w118 control line, F=Flag purification,
377 1/2= replicates and +/- = presence or absence of thiocyanate.
378
379 **Supplementary Table 2.** A snapshot of some of the proteins identified, numbers of
380 contributing peptides and % protein sequence coverage in Vha55 positive control and
381 corresponding W118 negative control experiments with (+) or without (-) thiocyanate.
382 Proteins highlighted in orange are either direct or indirect interactors of the bait
383 (green) and proteins highlighted in yellow are putative direct interactors based on
384 known interactions of orthologous proteins from other species. Proteins highlighted in
385 red are known contaminants based on previous experiments but those asterisked are
386 putative binding partners from yeast predictions. The remaining proteins (148) that
387 have no published interaction data have been excluded from the list. Green values
388 indicate increase of peptides or % sequence coverage and red highlights a decrease (or
389 increase where the 20 known contaminants are being measured).
390
391 **Supplementary Table 3. A.** Mass Spectrometry data from eight baits and
392 corresponding non-tagged controls to determine proteins being eluted non-specifically
393 from FLAG and Strep resins. **B.** Heat Map showing Mass Spectrometry data averaged
394 from all eight baits and corresponding non-tagged controls to demonstrate changes in
395 the average numbers of peptides and % sequence coverage after thiocyanate

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

396 treatment. Also shown are the Mascot scores for the baits to demonstrate that this is
397 not a good measure to show the effect of thiocyanate treatment.
398
399 **Supplementary Table 4.** Interaction lists and validation. **A.** Interaction lists for the 8
400 baits used to show proteins identified in FLAG and STREP affinity purifications with
401 or without the presence of thiocyanate ranked in decreasing order of numbers of
402 unique peptides, then % sequence coverage. * indicates the bait protein. Respective
403 *w118* negative control proteins found in FLAG and Strep pulldowns have been
404 removed but common 'known' contaminants have not. Contaminants identified in
405 pooled analysis are in italics. Grey highlighting indicates protein was detected with
406 and without thiocyanate. Grey boxes with white text identify proteins that have
407 interactions with the bait from orthologous species. Black highlighted boxes with
408 white text indicate interacting proteins unique to the inclusion or exclusion of
409 thiocyanate. **B.** Numbers of published and predicted interactors from DroID and the
410 numbers in our lists that were published using the FlyMine interaction queries.
411
412 **Acknowledgements.**
413 JSR wrote the manuscript. Both JSR and KSL conceived the ideas. This work was
414 funded by the Wellcome Trust grant 076739/Z/05/Z.
415
416 **Role of the funding source.**
417 The funding source, the Wellcome Trust, had no involvement in the design,
418 collection, analysis, interpretation of the data, in the writing of the report, or the
419 decision to submit the paper for publication
420
421 **References.**
422 1. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, et al. (2002) Functional
423 organization of the yeast proteome by systematic analysis of protein
424 complexes. Nature 415: 141-147.
425 2. Rees JS, Lowe, N. Armean, I.M., Roote, J., Johnson, G., Ryder, E., Russell, S., St
426 Johnston, D. and Lilley, K.S. (2011) In vivo analysis of proteomes and
427 interactomes using parallel affinity capture (iPAC) coupled to mass
428 spectrometry. Mol Cell Proteomics doi:10.1074/mcp.M110.002386.
429 3. Ranish JA, Brand M, Aebersold R (2007) Using stable isotope tagging and mass
430 spectrometry to characterize protein complexes and to detect changes in their
431 composition. Methods Mol Biol 359: 17-35.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

432 4. Wepf A, Glatter T, Schmidt A, Aebersold R, Gstaiger M (2009) Quantitative
433 interaction proteomics using mass spectrometry. Nat Methods 6: 203-205.
434 5. Trinkle-Mulcahy L, Boulon S, Lam YW, Urcia R, Boisvert FM, et al. (2008)
435 Identifying specific protein interaction partners using quantitative mass
436 spectrometry and bead proteomes. J Cell Biol 183: 223-239.
437 6. Kaake RM, Wang X, Huang L (2010) Profiling of protein interaction networks of
438 protein complexes using affinity purification and quantitative mass
439 spectrometry. Mol Cell Proteomics 9: 1650-1665.
440 7. Baldwin RL (1996) How Hofmeister ion interactions affect protein stability.
441 Biophys J 71: 2056-2063.
442 8. Zhang Y, Cremer PS (2006) Interactions between macromolecules and ions: The
443 Hofmeister series. Curr Opin Chem Biol 10: 658-663.
444 9. Richens JL, Lunt EA, Sanger D, McKenzie G, O'Shea P (2009) Avoiding
445 nonspecific interactions in studies of the plasma proteome: practical solutions
446 to prevention of nonspecific interactions for label-free detection of low-
447 abundance plasma proteins. J Proteome Res 8: 5103-5110.
448 10. Veraksa A, Bauer A, Artavanis-Tsakonas S (2005) Analyzing protein complexes
449 in Drosophila with tandem affinity purification-mass spectrometry. Dev Dyn
450 232: 827-834.
451
452

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

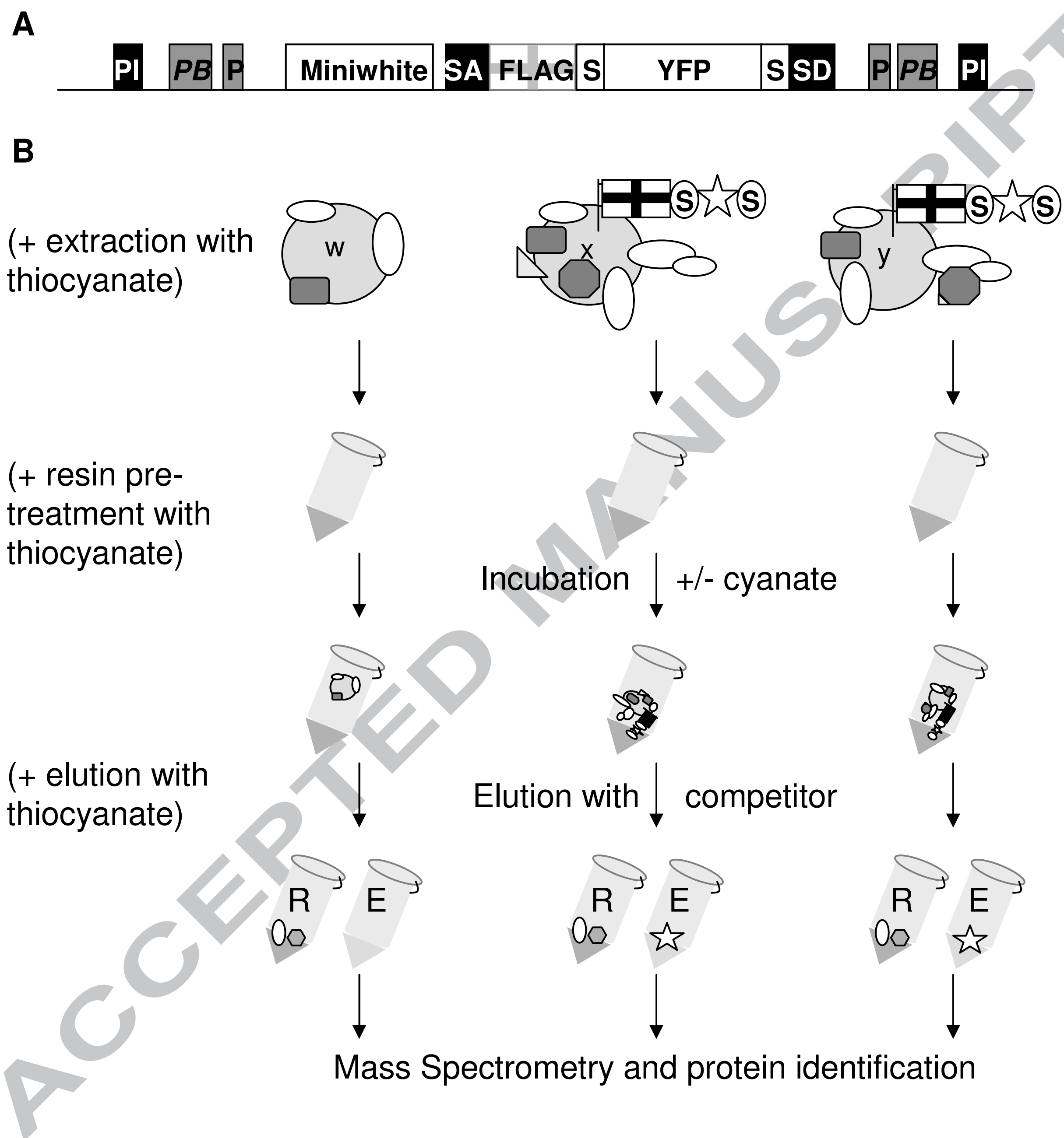

Figure 1. A. Construct used to generate tagged endogenous proteins. PI= Plasmid region, PB *PiggyBac* Inverted repeats, P= P-ends, SA= Splice Acceptor, FLAG= FLAG tag, S= *Strep*II tags, SD= Splice Donor. **B.** Workflow for the purification of *Drosophila melanogaster* tagged proteins with and without the inclusion thiocyanate at various steps indicated. S= Strep tag, w= w118 control line, x and y are tagged test lines. R= resin and e= eluate containing purified protein.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

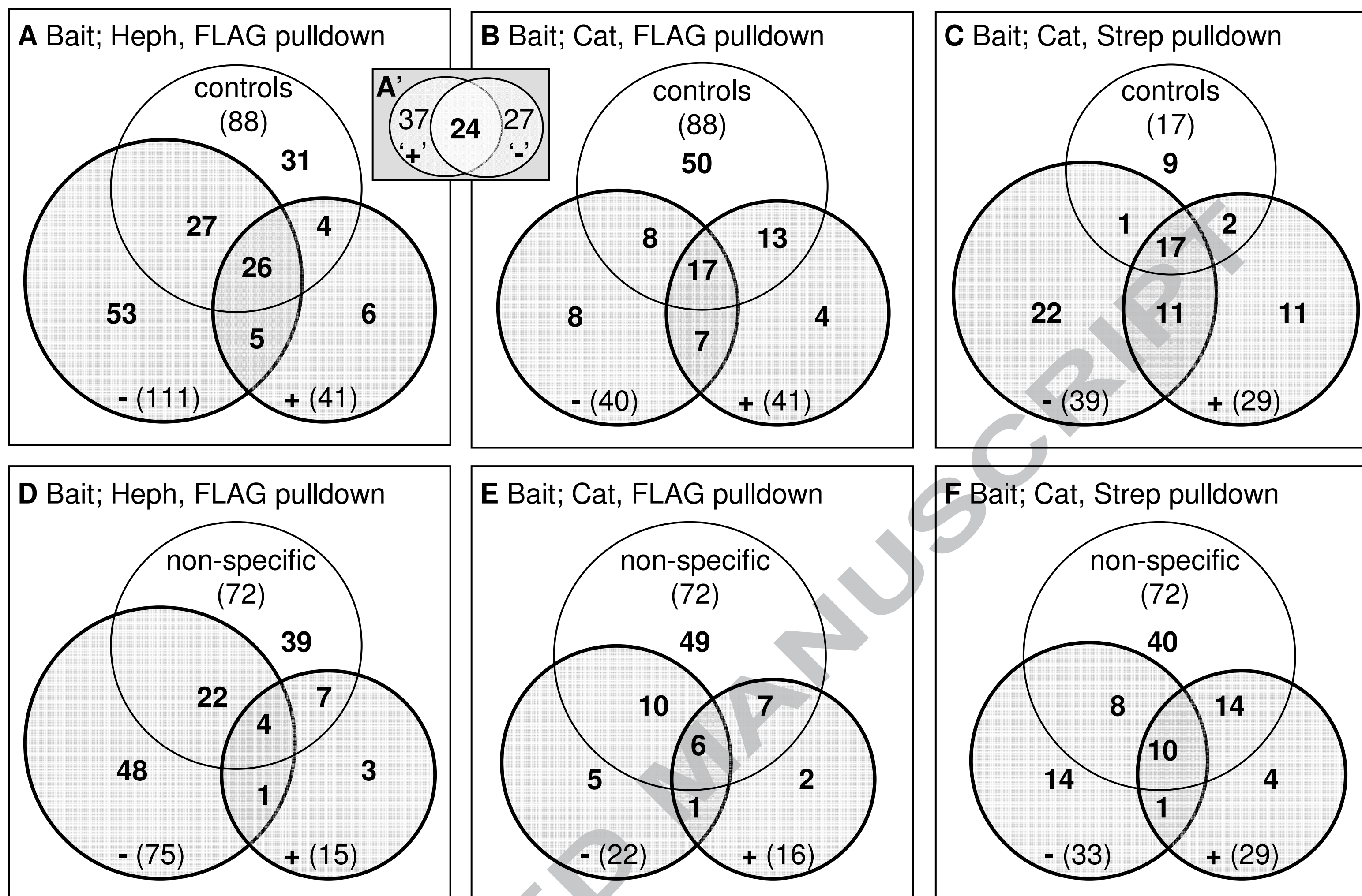

Figure 2. The effect of thiocyanate in reducing non specific protein binding to Flag and Strep affinity resins. **A-C.** Venn diagrams to show control proteins (Wf+&-) and eluted proteins from pulldowns in the presence (+) or absence (-) of thiocyanate. Subset Venn in **A'** is the merged negative control data. **D-F.** Venn diagrams showing each bait with its respective negative control data subtracted and the proportion of the data that is present more than 20% in all samples analysed 'non-specific' (list of proteins in Supplementary Table 3A).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

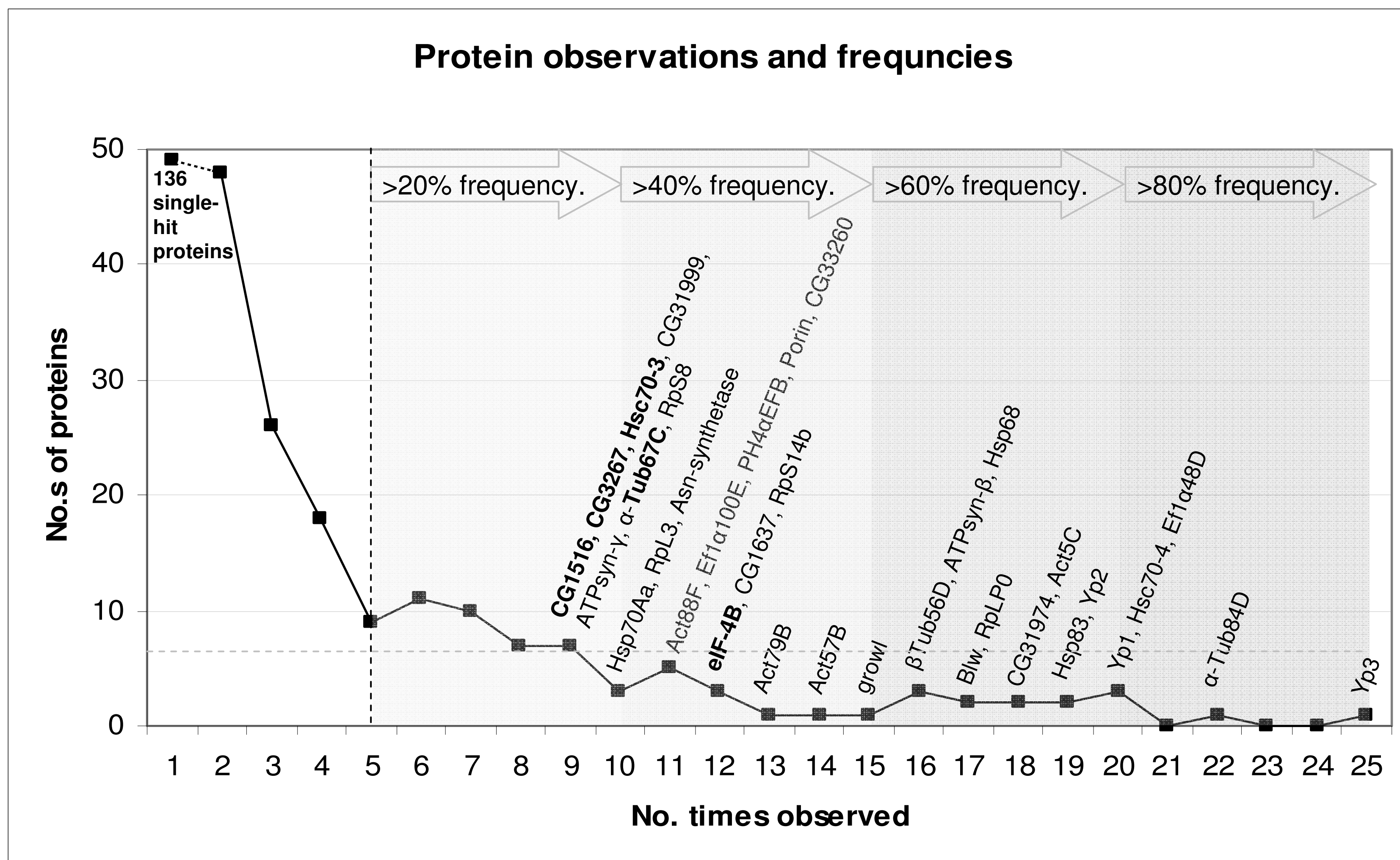

Figure 3A. The numbers, and some identifications of proteins identified in 25 experiments using 8 protein baits. Single hit proteins have been excluded. Shaded blocks indicate proteins occurring in >20%-100% frequently occurring interaction lists from all pulldown experiments (including negative controls) thus defined as 'contaminants'. These proteins are listed and detailed in Supplementary Table 3B. Dashed line shows the threshold we define our cut-off for contaminants. Proteins in bold were not frequently occurring in negative control samples so are likely to be bait specific and not resin specific (from Supplementary Table 1).

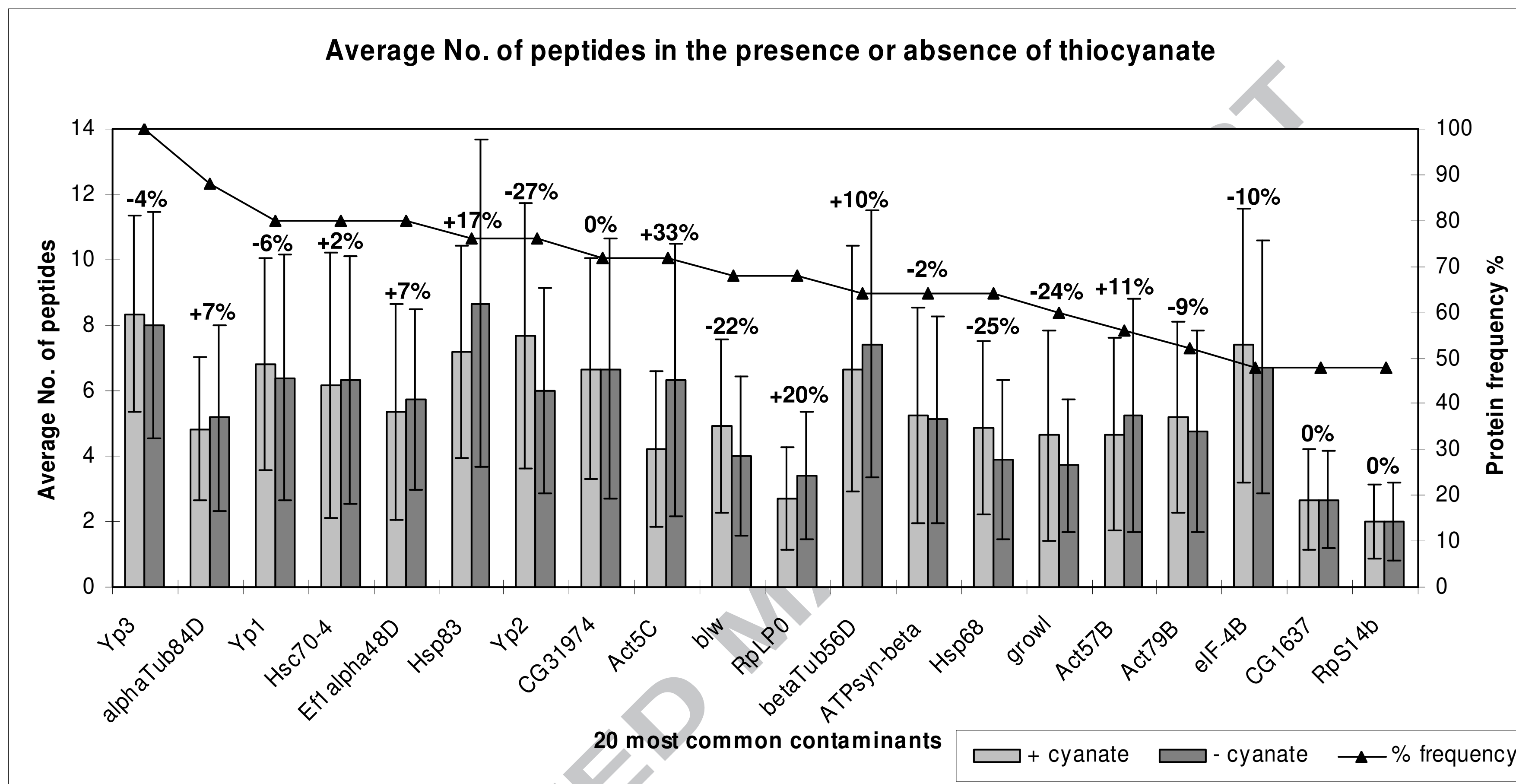

Figure 3B. An analysis of the numbers of peptides generated from the 20 most abundant proteins (with % frequencies in all interaction lists) observed in Figure 3A in the presence (light grey bars) or absence (dark grey bars) of thiocyanate with error bars. The most notable differences are in the scaffold proteins, actins and tubulins. The % average peptide count changes, compared with no treatment, are displayed above the bars.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Bait ID	Name/ FlyBase ID	size (KDa)	type	resin	bait peptides '+'	bait peptides '-'	bait seq. coverage (%) '+'	bait seq. coverage (%) '-'	YFP peptides '+'	YFP peptides '-'	YFP seq. coverage (%) '+'	YFP seq. coverage (%) '-'	Mascot score '+'	Mascot score '-'	% increase in bait peptides	% increase in bait coverage
2584	heph -PG	66.3	c	FLAG	7	6	21	18	12	12	30	41	390	106	16.67	16.67
	heph -PC	62.6	c	FLAG	6	0	17	0	12	12	30	41	375	0	>100	16.67
2785	CG11030	37.3	unk	FLAG	5	3	23	6	7	7	13	27	130	63	66.67	283.33
2786	Cat	57.1	m	FLAG	27	32	38	42	10	4	36	15	1066	1585	-15.63	-9.52
2796	shep	40-60	c	FLAG	3	2	8	11	7	8	20	25	57	75	50.00	-27.27
2818	Pop2	33.5	n&c	FLAG	6	3	28	10	7	6	28	18	89	80	100.00	180.00
3424	Aats-His	57/62	c	FLAG	9	6	19	14	4	3	10	13	366	407	50.00	35.71
2267	CG1440	55.2	c	Strep	47	39	62	48	4	6	21	42	982	853	20.51	29.17
2728	CG6084	36/40	c	Strep	68	42	66	43	36	36	38	33	898	1549	61.90	53.49
2786	Cat	57.1	m	Strep	34	24	43	39	10	5	31	21	1576	1124	41.67	10.26
2818	Pop2	33.5	n&c	Strep	0	0	0	0	5	2	34	6	0	0	0.00	0.00

Table 1. Mascot data from Mass Spec analysis of 8 tagged (bait) proteins purified in the presence (+) or absence (-) of thiocyanate. All FDRs are below 5%. C=cytoplasmic, m=membrane, n=nuclear and unk=unknown. PG/PC are distinct protein isoforms G and C respectively.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Large scale affinity purification studies reveal co-purifying contaminants.

Sodium thiocyanate (NaSCN) ions can help minimize persistent contaminants.

Importantly NaSCN also enriches for desired proteins and specific binding partners.

ACCEPTED MANUSCRIPT