Rapid Selection of Cyclic Peptides that Reduce Alpha-Synuclein Toxicity in Yeast and Animal Models

Rapid Selection of Cyclic Peptides that Reduce Alpha-Synuclein Toxicity in Yeast and Animal Models

Joshua A. Kritzer[1], Shusei Hamamichi[3], J. Michael McCaffery[4], Sandro Santagata[1,5], Todd A. Naumann[6], Kim A. Caldwell[3], Guy A. Caldwell[3] and Susan Lindquist[1,2]

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge MA 02142

[2]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139

[3]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487

[4]Integrated Imaging Center and Department of Biology, Johns Hopkins University, Baltimore, MD 21218

[5]Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA, and Harvard Medical School, Boston, Massachusetts, USA

[6]Department of Chemistry, The Pennsylvania State University, University Park, PA 16802

ABSTRACT

Display technologies have demonstrated the utility of cyclic peptides as protein ligands, but cannot access proteins inside eukaryotic cells. Expanding a novel chemical genetics tool, we describe the first expressed library of head-to-tail cyclic peptides in yeast. We applied the library to selections in a yeast synucleinopathy model that recapitulates much of the cellular pathology of Parkinson's disease. From a pool of five million transformants, we isolated two related cyclic peptide constructs which specifically reduce the toxicity of human α-synuclein. These expressed cyclic peptides also prevent dopaminergic neuron loss in an established *C. elegans* Parkinson's model. This work highlights the speed and efficiency of using libraries of expressed cyclic peptides for forward chemical genetics in cellular models of human disease.

Cyclic peptides (CPs) and their derivatives are potent bioactive compounds, and represent an underexplored, natural-product-like chemical space[1,2]. Phage and RNA display have made it possible to screen large libraries of CPs, cyclized via disulfide bonds or other side chain linkages, to identify high-affinity ligands for nearly any *in vitro* target[3,4]. By contrast, to date there are few methods for directly screening large libraries of CPs inside eukaryotic cells. Such methods would provide several advantages over *in vitro* techniques, ensuring that hits are nontoxic, can bind their target(s) in the appropriate cellular environment, are not rapidly degraded by cellular proteases, and possess at least enough selectivity to function in living cells. In addition, *in vivo* methods would enable

phenotypic selections of CPs, providing a more streamlined and less expensive alternative to traditional high-throughput screening with a greater chance of identifying effectors with non-traditional modes of action such as inhibition of protein-protein interactions.

Recently, Benkovic and colleagues reported a promising method of generating libraries of head-to-tail CPs *in vivo* using a single genetic construct named SICLOPPS (split-intein-mediated circular ligation of proteins and peptides, Fig. 1A)[5-7]. This construct uses a cleverly arranged split intein that splices out a linker region as a CP post-translationally; the linker can be as small as four amino acids or as large as a whole protein[8]. SICLOPPS-based CP libraries represent a powerful opportunity for rapid forward and reverse chemical genetics using *in vivo* selections[7]. Previous work by Benkovic and colleagues interfaced expressed CP libraries with bacterial two-hybrid selections, an elegant strategy for reverse chemical genetics[7]. Phenotypic screening of CP libraries was also performed in bacteria[9,10]. Despite these successes, to date there has been only one reported attempt to adapt SICLOPPS libraries to a eukaryotic system. A retroviral CP library was applied to a selection for inhibitors of interleukin-4 signaling in human B cells, yielding roughly a dozen CP pentamers with varying activity but no sequence consensus[11]. The general utility of this library was hampered by its low actual diversity (2.7×10^5 members), the lack of quantitative quality assessment, and the several weeks required just to perform an initial round of selection.

We sought to apply expressed CP libraries to eukaryotic cells and perform phenotypic selections in models of human disease in a rapid, efficient and generally applicable manner. This goal dovetailed with our ongoing investigations of protein misfolding diseases, since we and others have postulated that targeting such diseases will require non-

traditional modalities such as inhibition of protein-protein interactions or modulation of the proteostasis machinery[12-14]. We therefore turned to CP libraries as a source of molecules most likely to act via these mechanisms, and to yeast models of protein misfolding diseases as testing platforms for this novel approach.

We and others have demonstrated that, because protein misfolding often affects highly conserved biological pathways, complex diseases such as Parkinson's disease (PD) can be modeled in simple organisms such as yeast[15-20]. The human protein α-synuclein (α-syn) has been linked to Parkinson's disease (PD) via genetic evidence and its prominence in the PD-associated intracellular aggregates known as Lewy bodies[21-23]. α-syn is a small lipid-binding protein that is prone to misfolding and aggregation, and in the yeast *Saccharomyces cerevisiae* expression of human α-syn over a threshold level leads to ER stress, disruption of ER-Golgi vesicle trafficking, accumulation of lipid droplets, mitochondrial dysfunction, and ultimately cell death[18-20]. This cellular pathology mirrors many aspects of dysfunction seen in neurons and glia of individuals with PD and other synucleinopathies[23]. Genetic screening using our yeast synucleinopathy model has yielded suppressors of α-syn toxicity that are also effective in neuronal models[19,20,24]. Moreover, genetic analyses in these models have directly linked α-syn toxicity to the function of *PARK9*/ATP13A2, a protein whose mutations lead to an early-onset form of PD but otherwise had no known connection to α-syn[24,25]. Thus, cellular models have been critical to our understanding of α-syn and its role in the selective degeneration of dopaminergic (DA) neurons in PD. However, despite these and other intensive efforts, there remains a paucity of proven pharmacological targets for PD and other synucleinopathies.

Taking advantage of our established yeast synucleinopathy model, we constructed the first yeast-compatible CP library and used rapid phenotypic selections to isolate CPs that specifically reduce α-syn toxicity. Further, we identified the CP motif responsible for activity and demonstrated that the selected CPs significantly reduce DA neuron loss in a *C. elegans* PD model. These advances establish *in vivo* CP selections as an immediately useful tool for a broad range of biologists.

RESULTS

The SICLOPPS construct expresses and splices in yeast

We obtained a SICLOPPS construct (a generous gift from Dr. Stephen Benkovic) and expressed it on a high-copy 2μ plasmid in *Saccharomyces cerevisiae* under a constitutive *PGK* promoter. Lysates from log-phase and stationary-phase cultures were blotted using an antibody to a C-terminal tag (Fig. 1B), which revealed the expression level and splice status of the split intein construct[8]. The log-phase cultures showed approximately equal amounts of 26 kDa and 20 kDa proteins, which are the expected sizes of the unspliced and spliced constructs, respectively. Cultures harvested in stationary phase showed only the 20 kDa band corresponding to the spliced form. These findings are consistent with a splicing reaction that is somewhat slower than the rate of log-phase *PGK*-driven protein synthesis.

A splice-disabled version of the gene was constructed by mutating the residues corresponding to T69 and H72 of the *dnaE* split intein to alanine. Previous work has established that these mutations, which reside in conserved block B, permit association of

the two halves of the split intein but prohibit processing at the extein/N-intein junction. This blocks the first obligatory step in the accepted splicing mechanism[26-28]. When a T69A/H72A mutant of the HPQ construct was expressed in yeast, only the full-length, unspliced form was detected (Fig. 1B). Thus, the lower 20 kDa band indeed represents the expected product of intein splicing.

To verify directly that the SICLOPPS construct produces CPs in yeast, we purified the encoded CP from yeast lysates by virtue of its HPQ motif. This sequence has well-established streptavidin affinity which was previously used to purify spliced CPs from bacterial cultures[29]. We used streptavidin agarose to pull down the HPQ CP from lysates of stationary-phase yeast and confirmed its identity by MALDI and electrospray mass spectrometry ($MW_{(M+H)}$=1330.6, MW_{obs}=1330.5). While the native yeast *Vma1* intein has long been known and studied, these data provide the first demonstration that the SICLOPPS construct can be used to produce CPs in yeast.

Construction of a yeast-compatible CP library

Having established that the SICLOPPS construct expresses and splices in yeast, we sought to construct a high-diversity library compatible with yeast-based selections. We used a previously described PCR cloning strategy consisting of PCR with a randomized oligonucleotide followed by a second PCR with constant primers to generate a randomized insert with a minimum of frameshift errors[30]. In contrast to previous SICLOPPS libraries with smaller CP backbones, we chose to construct an octamer library to maximize diversity and minimize effects of side chains on splicing efficiency[6,8,11]. The library was designed to encode CP octamers with a single cysteine residue followed by

seven randomized residues, since a nucleophile at the first position is required for the transesterification step of intein processing [26]. DNA was randomized using the NNS codon scheme, where N is any nucleotide and S is guanine or cytosine. This allows representation of all twenty amino acids while reducing the impact of degeneracy and stop codons. After transformation into bacterial cells, the total diversity of the library (as measured by counting the number of independent transformants by serial dilution) was 5 x 10^7 independent members.

Library quality was assessed by picking and analyzing twenty colonies at random (Fig. 1C and supplement). Sequencing showed that library constructs encoded CPs with high diversity at the protein sequence level. Sequence data also confirmed that the library possessed low background: only one of twenty clones was the original HPQ-encoding vector, three of twenty possessed frameshift mutations, and one of twenty had a stop codon in the randomized region. Thus, ~75% (15 of 20) of the library is estimated to encode novel CPs.

We next assessed expression and splicing of each of the twenty randomly picked constructs. The laboratory yeast strain W303 was individually transformed with each of the twenty constructs, grown to log phase, and lysed. Western blots of these lysates revealed that all fifteen of the novel CP constructs expressed. Fourteen showed evidence of some splicing at log phase, most of them quite robustly (Fig. 1C). This quality control analysis indicates that roughly 70% of the library encodes CP constructs that express and splice in yeast, representing an overall real diversity of over thirty million CPs. The library at least 0.1% of the overall theoretical diversity, and it represents a vast, diverse, and largely unexplored chemical space ready for screening or selection.

Selection in a yeast synucleinopathy model

With a diverse, high-quality library in hand, we next sought to perform selections in a complex biological system, one which would not be amenable to *in vitro* methods such as phage display. To this end, the library was applied to our yeast synucleinopathy model using a simple bulk selection protocol. The yeast synucleinopathy model uses the *GAL1* promoter to tightly control expression of human α-syn, so that no α-syn is produced until the yeast are grown in galactose media[18,19]. One liter of yeast culture was transformed with 80 μg library DNA and plated onto large library plates lacking uracil to select for transformed cells. This yielded five million independent transformants, which were scraped, re-plated onto galactose media to induce α-syn expression, and incubated for three days (Fig. 2A). Ninety-six large colonies were picked and the corresponding plasmids were isolated, amplified in bacteria, and tested by individually re-transforming each into the screening strain. Thirty-one clones demonstrated reproducible suppression of α-syn toxicity in yeast, four of which are shown in Fig. 2B. False positives likely stemmed from spontaneous genomic suppressor mutations, which are common in most yeast selections[19,31].

Filtering assays

The thirty-one selected CP constructs were further evaluated using two secondary assays to exclude promoter effects and other nonspecific modes of action. First, the constructs were transformed into a reporter yeast strain with the *LacZ* gene downstream of the *GAL1* promoter[32]. Expression of β-galactosidase from the *GAL1* promoter was then

quantified using the soluble β-galactosidase substrate chlorophenol red-beta-D-galactopyranoside (CPRG). This enabled us to quantify the effects of each construct on *GAL1*-mediated expression independently of the presence of α-syn (Fig. 2C). Second, the constructs were transformed into a yeast disease model that uses the *GAL1* promoter for the inducible expression of a toxic polypeptide unrelated to α-syn, exon 1 of human huntingtin with 103 glutamine residues (Htt-103Q)[16,17]. The transformants were then spotted onto galactose media to assess the abilities of selected hits to suppress the toxicity of Htt-103Q in yeast (Fig. 2D). This assay provided a second check on whether the constructs acted by interfering with expression from the *GAL1* promoter, and also tested whether the observed effects were specific to α-syn. Of course, selected CP constructs that suppressed both α-syn and Htt-103Q toxicity could be affecting a pathway common to both[33]. However, all selected constructs that prevented Htt-103Q toxicity also reduced *GAL1*-mediated expression (for example, CP3 shown in Fig. 2B-D). Two of the thirty-one selected constructs, denoted CP1 and CP2, suppressed α-syn toxicity but had no effects in the *GAL1* reporter assay or on Htt-103Q toxicity, and thus act via an α-syn-specific mechanism. Notably, these results are in keeping with our previous results from genetic screens, in which genes that suppress α-syn toxicity possessed little overlap with genes that suppress Htt-103Q toxicity[17,19]. Thus, suppressors such as CP1 and CP2 affect specific pathways of eukaryotic cell biology involved in the toxicity of these human proteins, rather than nonspecific pathways involved in general protection from misfolded or aggregated proteins.

At this point we sought to ensure that the spliced CPs were responsible for the observed suppression of α-syn toxicity rather than nucleic acid or peptide aptamers

encoded by the selected constructs. To this end, we performed three independent tests to ensure complete splicing was required. We first tested T69A/H72A mutants of each construct, which are unable to process at the extein/N-intein junction[26,27]. T69A/H72A mutants of CP1 and CP2 were unable to suppress α-syn toxicity (Fig. 2E). Next, a different set of mutations at residues H24 and F26 in the *dnaE* C-intein domain (responsible for promoting asparagine cyclization in the final processing step at the extein/C-intein junction)[26,34] were also shown to abolish the activities of CP1 and CP2 (see supplement). Thus, the complete splicing mechanism, including processing at both the N-intein and C-intein junctions, is required for CP1 and CP2 activity. Finally, to completely rule out the possibility that linear peptide byproducts are responsible for suppression of α-syn toxicity, we also directly tested alternative constructs that represent products of an N-intein cleavage event rather than complete cyclization. These constructs also failed to suppress α-syn toxicity, demonstrating that the sequences encoded by CP1 and CP2 are inactive without complete intein splicing (see supplement for details). Remarkably, one of the strongest inhibitors of expression from the *GAL1* promoter, CP3, also required splicing for function, demonstrating that its mode of action is also mediated by a spliced CP.

Structure-activity relationships by point mutagenesis

A critical bottleneck in chemical genetics is the generation of useful structure-activity relationship (SAR) data once hits are identified from large compound libraries. SAR data are essential for understanding a molecule's mode of action, for the development of more potent derivatives from initial hits, and for the derivatization of molecules as affinity

reagents. A great advantage of the CP approach is that SAR data can be generated rapidly and cost-effectively using point mutagenesis. The sequences of the CPs encoded by CP1 and CP2 are shown in Figure 3A. We generated a series of alanine-scanning mutants to identify which side-chains of CP1 and CP2 are responsible for function, and to determine whether the two suppressors share a common functional motif. Each mutant was transformed into yeast and expression and splicing were verified by blotting log-phase cultures. Yeast cells expressing each mutant were serially diluted and plated on galactose media to quantify the degree of α-syn toxicity suppression (Fig. 3B). We found that the cysteine in the fourth CP position was absolutely required for activity in both constructs. Mutation of the hydrophobic residue in the third CP position (tryptophan for CP1 and leucine for CP2) resulted in constructs with very weak activity. Mutation of the residue in the fifth CP position (serine for CP1 and glutamate for CP2) resulted in a slight but reproducible decrease (~5-fold) in growth, indicating a possible role for these side chains as well.

Replacement of the cysteine in the first position with alanine resulted in another splice-disabled construct, rendering the mutant construct inactive (Fig. 3B). To test the requirement for the first cysteine thiol in the spliced CP, we mutated it to serine. These mutants expressed and spliced as efficiently as the original CP1 and CP2 constructs as assessed by blotting, but did not suppress α-syn toxicity (Fig. 3B). Thus the cysteine in the first position is also required for the biological activities of selected CPs. Overall, CP1 and CP2 share a common CXΦC motif, wherein X is any residue and Φ is a hydrophobic residue; their functional motifs are highlighted in Figure 3A.

Minimization of the CP backbone using iterative design

We constructed an octamer CP library to provide diversity, but it was unclear whether selected hits would be functional only as octamers or might be amenable to minimization. Minimization of CPs would evaluate the extent to which selected functional motifs are dependent on macrocycle size and structure, and would also be desirable for downstream applications as chemical genetics agents. Having selected two functional CP constructs with a common tetrapeptide motif, we had the opportunity to test directly whether the remaining portions were necessary for proper orientation of the motif. Thus, building upon our side chain SAR data, we designed further mutants in order to minimize CP ring size.

CP1 proved especially amenable to minimization, as shown in Figure 3C. Residue 8 could be deleted entirely with no loss of function. However, a construct with both residues 8 and 7 deleted showed only minimal function and a construct with residues 8, 7, and 6 deleted showed no ability to rescue α-syn toxicity. Since the side chains of these residues are not required for function, we hypothesized that conformational effects on the CXΦC motif were responsible for the loss of function. To restore proper orientation in the context of the hexamer, we tested a variety of substitutions in positions 2, 5 and 6, including glycine, alanine, proline, lysine, and tryptophan. Of the twelve mutants tested, substitution of the threonine at position 6 with alanine or glycine had the greatest effect, restoring the ability to suppress α-syn toxicity to an extent comparable to that of the original eight-residue CP1 (Fig. 3C). Splice-disabled mutants of these constructs were also tested to ensure that activity was still dependent on intein splicing. Overall, two rounds of rationally designed mutagenesis generated constructs encoding CP hexamers

that suppress α-syn toxicity equally as well as the selected CP octamers CP1 and CP2. Thus, simple iterative testing of designed mutants was used to minimize the originally selected eight-residue CP (predicted MW=933) to a six-residue CP (predicted MW=634) with no observable change in *in vivo* activity.

CP1 and CP2 are easily incorporated into affinity reagents

Target identification is another notoriously difficult hurdle in forward chemical genetics[35]. CPs are synthetically accessible in multi-milligram quantities and easily derivatized, making them ideal for physical pull-down experiments. After verifying genetically that incorporation of a lysine at position 7 did not affect CP function (Fig. 3C), we synthesized $CP1_{R7K}$ and $CP2_{W7K}$ (see supplement). Since the CPs have no free N-termini, the lysine amines represent useful handles for site-specific derivatization or attachment to solid phase. We linked the synthetic CPs to agarose beads and optimized pulldown conditions. Notably, high lysate protein concentrations (5-10 mg/mL) could be used without extensive nonspecific binding due to the hydrophilic nature of the CPs.

Proteins that selectively bound CP1 or CP2 beads over control CP-linked beads were eluted, separated by SDS-PAGE, and identified by mass spectrometry. Yeast genetics greatly streamlines the process of validating potential targets, and so we were able to rapidly test several candidates including the chaperone proteins Ssa1 and Ssc1. However, modulation of the levels of these proteins in yeast by genetic overexpression or deletion did not affect α-syn toxicity or CP-mediated suppression of α-syn toxicity. We also easily linked $CP1_{R7K}$ and $CP2_{W7K}$ to biotin and to photoactivatable cross-linking groups, but were unable to selectively isolate additional target candidates. There are three likely

explanations for these results. First, CP targets may be transient, unstable, or present at low levels within the cell. Second, selected CPs may possess low affinities for their targets, as CPs selected in bacteria showed high micromolar K_i's and IC_{50}'s in *in vitro* assays[6,36]. Finally, CP1 and CP2 may act via a nontraditional mechanism that is incompatible with physical target identification (see discussion). Our results demonstrate that the preparation of CP-based affinity reagents can be rapid and straightforward. We are currently exploring further methods to translate this capability into similarly rapid techniques for target identification.

CP1 and CP2 act in a manner that is distinct from known suppressors of α-syn toxicity

Previous work using the yeast model has established that overexpression of human α-syn affects multiple steps in vesicle trafficking[20]. We wondered whether our selected CPs act upstream or downstream of this defect. The yeast synucleinopathy model is particularly useful for examining the effects of α-syn on vesicle trafficking because α-syn induction can be tightly controlled and α-syn localization within the cell can be monitored in real time using YFP fusions[19,20]. Shortly after induction, α-syn-YFP accumulates on the plasma membrane. After 2-3 hours of α-syn-YFP induction, small foci are observed peripherally to the membrane, and after 4-6 hours these foci coalesce and begin to migrate inward toward the vacuole. This latter stage is associated with the onset of cell death. Immunoelectron microscopy has revealed that these foci correspond to pools of stalled vesicles that are coated with α-syn-YFP[20].

We wondered whether CP1 and CP2 prevent the accumulation of pools of stalled vesicles, since previously described genetic suppressors of α-syn toxicity mitigate the trafficking blockage[19]. For instance, the expression of a Rab homologue involved in ER-Golgi trafficking (Ypt1) or the yeast homologue of the P-type ATPase *PARK9*/ATP13A2 (Ypk9) suppresses α-syn-YFP foci (Fig. 4A)[19,20,24]. By contrast, CP1 and CP2 do not alter the occurrence or appearance of α-syn-YFP foci after 4 hours. Electron microscopy confirms that cells co-expressing CP1 or CP2 along with α-syn still possess pools of stalled vesicles similar to those found in cells expressing a negative control CP along with α-syn (Fig. 4B). Remarkably, even though 8-12 hours of α-syn-YFP induction is lethal to control cells, CP1 and CP2 possess α-syn-YFP foci continuously for over 48 hours after induction, with little change in appearance after 8 hours (Fig. 4A). Thus, CP1 and CP2 permit cells to grow and divide, albeit at a reduced rate, despite the continuous presence of stalled α-syn-coated vesicles. This phenotype is unique among the genetic and small molecule suppressors of α-syn toxicity identified to date, and implies that selected CPs target a pathway downstream of established vesicle trafficking defects[19,20,24,33].

The previously described genetic suppressors Ypt1 and Ypk9 act in independent pathways, so that the simultaneous expression of both genes results in more potent suppression of α-syn toxicity than the expression of either gene by itself[24]. To genetically test whether selected CPs act in a novel pathway, we expressed them in combination with Ypt1 or Ypk9. The effects of CP1 and CP2 combine in an additive manner with the effects of either of these genetic suppressors (see supplement). Thus, the CPs likely act in pathways that are distinct from those already linked to α-syn toxicity.

Expression of CPs in a metazoan animal model

We previously demonstrated that human homologues of suppressors of yeast α-syn toxicity identified were able to rescue dopaminergic (DA) neurons in animal models of synucleinopathy[19,20]. In a *Caenorhabditis elegans* model, expression of human α-syn from the dopamine transporter promoter P_{dat-1} results in degeneration of DA neurons[20,37,38]. This enables direct evaluation of putative α-syn antagonists in DA neurons in the context of whole animals. Nematode development is highly stereotyped and reproducible, so that any alteration in neuron quantity or morphology is highly significant. Nematodes are also transparent, enabling assessment and scoring of DA neurons in live animals by virtue of GFP expression under the control of the *dat-1* promoter. Using this model, we have found that expression of genes such as Rab1 and *PARK9*/ATP13A2 (homologues of Ypt1 and Ypk9) ameliorate α-syn toxicity in the DA neurons of live animals[20,24].

We sought to evaluate whether CP1 and CP2 would similarly rescue DA neurons in the nematode synucleinopathy model. Because the original split intein was isolated from a cyanobacterium, we first codon-optimized and re-synthesized the entire gene to ensure efficient expression in metazoan systems such as *C. elegans* and mammalian cells. Then P_{dat-1}*::mCP-gfp* expression vectors were constructed by recombinational cloning with the pDEST-DAT-1 destination vector to generate a construct that would express the mCP gene with a C-terminal GFP fusion in nematode DA neurons.

First, to verify that the mCP construct is expressed in the DA neurons of *C. elegans*, we injected P_{dat-1}::*mCP-gfp* expression vectors into wild-type animals. Because nematodes have only eight DA neurons, detection of gene products expressed exclusively

in these few cells by Western blotting is problematic. However, GFP fusions allow direct observation of protein expression in live animals. Expression of mCP-GFP started during embryogenesis, and in adult animals robust GFP fluorescence was observed selectively in DA neurons (see supplement). Because the GFP is encoded downstream of the mCP gene, we inferred that the mCP gene expresses in *C. elegans* DA neurons.

CP1 and CP2 rescue DA neurons in a *C. elegans* synucleinopathy model

We next tested the effects of expressed CP1 and CP2 on α-syn-mediated DA neuron loss in *C. elegans*. The CP1 and CP2 sequences were introduced into mCP genes without GFP fusions by site-directed mutagenesis and the resulting constructs were cloned into the pDEST-DAT-1 vector[19,20,24]. Transgenic worm lines were generated using these expression vectors as described[19,38]. We also constructed and tested negative controls, denoted CP1* and CP2*, with the splice-disabling T69A/H72A mutation.

We quantitated the effects of mCP constructs on α-syn toxicity as previously described for human and nematode genes, staging individual worms and inspecting the six anterior DA neurons (2 ADE and 4 CEP neurons)[19,20,38]. Representative fluorescence micrographs of worms and population phenotype analyses are presented in Figure 5. Expression of all mCP constructs had no discernable effect on neuron health or morphology in the absence of α-syn, indicating the mCP gene is well-tolerated by live animals in DA neurons. As previously described, GFP expression also had no effect on DA neurons, while α-syn expression caused striking DA neurodegeneration [19,20,24,38]. Only 15% of 7-day old worms expressing α-syn retained wild-type numbers and morphologies of DA neurons. Expression of mCP constructs encoding either CP1 or CP2

in these animals significantly reduced DA neurodegeneration, raising the proportions of animals devoid of neurodegeneration to 26% and 24%, respectively (Fig. 5G). This degree of rescue is comparable to that observed with human Rab1, Rab3a or *PARK9*/ATP13A2 in the same system[20,24]. Splice-disabled variants CP1* and CP2* were ineffective.

We also quantitated the severity of the neurodegeneration by counting the number of healthy DA neurons in each animal studied. We observed that 15% of animals expressing α-syn alone retained all six anterior DA neurons, while 40% retained five or more and 80% retained four or more (Fig. 5H). CP1 and CP2 significantly increased the proportion of worms with all six DA neurons intact and the proportion of animals that retained five or more DA neurons, consistent with our observations that worms expressing α-syn are most susceptible to degeneration in their two ADE neurons, and less susceptible in their four anterior CEP neurons[20,24,38]. Animals expressing splice-disabled variants CP1* and CP2* showed similar distributions to those expressing α-syn alone, again demonstrating that intein processing is required for the observed activities of the CP1 and CP2 constructs. Overall, the amelioration of α-syn toxicity afforded by CP1 and CP2 matches those of endogenous regulators of PD-associated cellular pathways such as Rab1, Rab3a or *PARK9*/ATP13A2[20,24].

DISCUSSION

We report a rapid and versatile technique for the *in vivo* selection of bioactive CPs in yeast. This technique represents a general chemical genetics strategy for identification of

functional modifiers of diverse cellular pathways. Our CP octamer library encodes diverse macrocycles of molecular weights from ~600 to 1200 daltons and can be effectively introduced into any yeast strain, including models of human disease. Hits can be rapidly identified using selections or screens that process millions of CPs in a single day without expensive robotics. The genetic tractability of yeast ensures that diverse filtering assays can be applied in high-throughput fashion to isolate the most relevant and potent hits. In selections using our yeast synucleinopathy model, two hits were isolated from an original pool of five million after only a single round of selection. While additional rounds of selection were not required for selections in the synucleinopathy model, we note that multiple rounds of selection could be performed by pooling colonies and amplifying their plasmids *en masse*. Stringency of the selection could be increased in subsequent rounds by using more robust selection conditions or by transferring selected CP genes to expression vectors with weaker promoters.

We also report the first expression of optimized mCP genetic constructs in live metazoan animals. Nematodes tolerate the construct well, with no observed effects on DA neuron health or morphology after lifelong expression of several different mCP constructs. As in the yeast model, splice-disabled mutants were used to verify that the spliced CP, and not a nucleic acid or protein aptamer, was responsible for reducing α-syn toxicity. The extents to which CP1 and CP2 reduce the amount and severity of DA neurodegeneration match those of human homologues of hits from our yeast genetic screen[19]. These include Rab GTPases critical for proper vesicle trafficking and *PARK9*/ATP13A2, a putative P-type ATPase linked to a genetic, early-onset form of

parkinsonism[19,20,24,25]. Thus, in yeast and *C. elegans* synucleinopathy models, selected CPs are as potent as endogenous regulators of key cellular pathways implicated in PD.

The ability to use mutagenesis to generate SAR data and the straightforward synthesis of CPs greatly streamlines the turnaround of hits into affinity reagents for target identification. The time from initial selections to having an affinity reagent in hand was less than two months. Physical pull-downs have not yet revealed a clear target for CP1 and CP2, but various alternatives are being employed to identify their modes of action. For instance, we are currently developing CPs with higher potency by designing more stringent selections and applying second-generation libraries that incorporate the CXΦC motif. We are also exploring whether CP1 and CP2 reduce α-syn toxicity by buffering the thiol redox state of the cytosol or by otherwise altering the activity of thioredoxin-fold proteins[39]. These possibilities are suggested by the striking similarity between the CXΦC motif and the thioredoxin-fold CXXC motif, the critical roles of thiol redox pathways in ER stress (which is a prominent effect of α-syn expression), and previous findings that dithiol small molecules, peptides and small CPs have biologically significant thiol redox potentials that vary greatly depending on neighboring functional groups and backbone conformation[39-42]. The CXΦC motif may also function by binding metal ions, since metal ion transporters were discovered among the genetic suppressors and enhancers of α-syn toxicity and manganese exposure has long been known as a risk factor for PD-like syndromes[24,43]. Either of these possibilities would represent novel proteostasis-based mechanisms that would defy physical target identification. Finally, a variety of genetic and genomic tools are available for target identification in yeast, including transcriptional

profiling, genetic screening, and the use of libraries of bar-coded overexpression and deletion strains[35,44]. All of these strategies will be applied to CP1 and CP2.

CPs selected in the yeast synucleinopathy model produce an unexpected and unique phenotype, apparently acting downstream of α-syn-mediated vesicle trafficking defects. Moreover, CP1 and CP2 act independently of known genetic suppressors of α-syn toxicity, and are effective in both yeast and nematode synucleinopathy models. This provides critical evidence that the target pathways are indeed conserved and relevant to the pathobiology of α-syn in neurons. Whether they act by binding a specific protein or by modulating thiol or metal homeostasis, CP1 and CP2 represent a promising new avenue for the exploration of therapeutics for PD and other synucleinopathies. Moving forward, the successful application of the new mCP construct in a metazoan animal opens the door to further validation in other PD models, including viral introduction into cell culture and incorporation in transgenic mice[20,24,45].

Overall, the CP approach represents an empowering technology for the rapid development of chemical genetics agents for systems that are difficult to address *in vitro*. In addition, the abundance of established yeast two-hybrid selection strains, already validated in explorations of the binary interactomes of yeast, *C. elegans* and man, will enable rapid and low-cost selections for CP inhibitors of specific protein-protein interactions at an unprecedented scale[46-48]. Our approach provides a powerful alternative to traditional small molecule screening, straying from traditional "drug-like" chemical space but offering vastly higher throughput and lower cost. It also possesses distinct advantages over display technologies, requiring fewer rounds of selection, obviating the need to purify and immobilize the target, and enabling selections *in vivo*, albeit with lower

throughput. These trade-offs will appeal to biologists interested primarily in developing tools for probing specific cellular pathways and identifying key druggable proteins in disease models.

Acknowledgments

We thank S. Benkovic for plasmids, E. Spooner for mass spectrometry assistance, and N. Azubuine for media preparation. This work was supported by an NRSA fellowship from NINDS/NIA (J.A.K.), an R21 grant from NINDS (S.L.L. and J.A.K.), and the Morris K. Udall Centers of Excellence for Parkinson's Disease Research (S.L.L.). Parkinson's disease research in The Caldwell Lab (G.A.C, K.A.C., S.H.) was supported by the American Parkinson Disease Association, Michael J. Fox Foundation, and NIEHS.

METHODS

Yeast strains, manipulation, selections and blotting

The α-syn-expressing yeast strains used in this study were all haploid W303 strains with the following genotype: *MATa can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 ade2-1*. Selections and SAR data generation was performed in the previously described IntTox strain, which has human α-syn integrated using *pRS303-αSynWT-YFP* and *pRS304-αSynWT-YFP* vectors[19]. Co-expression of CPs with genetic suppressors was performed using the related HiTox strain, in which human α-syn is integrated using *pRS304-αSynWT-GFP* and *pRS306-αSynWT-GFP* vectors[18]. The Htt-103Q strain used was

similarly constructed from the W303 background using integrated *pRS303-Htt103Q-GFP* and *pRS305-Htt103Q-GFP* constructs as described[16].

Yeast cultures were transformed using standard lithium acetate heat shock. In all cases, at least three independent transformants were isolated and tested to ensure reproducibility of the resulting phenotype. Library transformations were performed using a scaled-up heat shock procedure. One liter of IntTox yeast were grown to log phase in SC –his –trp media, pelleted and washed several times with water. Pellets were resuspended in 20 mL of 0.1 M lithium acetate and incubated for 20 minutes at 30 °C with shaking. Cells were then spun down and resuspended in 10 mLs of 0.1 M lithium acetate with 10 mM Tris-Cl, pH 7.5. 1 mM EDTA. 80 µg library DNA was added along with 1.4 mLs of carrier DNA (salmon sperm DNA, 2 mg/mL, previously boiled and kept on ice) and 25 mLs of 50% PEG-4000. After 30 minutes of incubation at 30 °C with shaking, 3 mLs of DMSO was added and the cells were heat shocked at 42 °C for 23 minutes. Cells were pelleted and resuspended in water prior to plating on large 24.5 x 24.5 cm Petri dishes of SC –his –trp –ura solid medium. After two days, all colonies were scraped and pooled, and 150,000-200,000 cells were plated on each of 35 SGal –his –trp –ura plates. Picked colonies were cultured overnight in SC –ura and pelleted. Plasmids were isolated using the PrepEase yeast plasmid isolation kit (USB).

Cultures for blotting were produced by growing cultures until saturation (overnight), diluting to an OD_{600} of 0.1 and growing for 4-5 hours. Lysates were produced by bead-beating pelleted cells in ethanol with 1 mM phenylmethylsulphonyl fluoride (PMSF), removing the ethanol by evaporation in a Speed-Vac (Savant), and resuspending in 20 mM Tris with 2% SDS. These were then normalized by protein concentration using a

BCA assay (Pierce), mixed with SDS-free loading buffer and subjected to SDS-PAGE and blotting onto PVDF membranes. Blotting was performed using anti-CBD serum (New England Biolabs) or anti-HA purified antibody (Sigma) according to standard protocols.

HPQ cyclic peptide isolation

Log-phase cultures expressing the HPQ-encoding SICLOPPS construct were spun down and resuspended in 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, with protease inhibitors (Roche) and 1 mM DTT. Cells were lysed by passing through a French press at 10,000 psi twice. Lysates were incubated with streptavidin-agarose beads for 4 hours at 4 °C and washed extensively. 1 mM biotin was used to elute the HPQ cyclic peptide. Eluates were freeze-dried using a Speed-Vac (Savant) and resuspended in 10 μL water/acetonitrile with 0.1% formic acid. These were further cleaned using C18 ZipTips (Millipore) and analyzed by MALDI-MS ($MW_{(M+H)}$=1330.6, MW_{obs}=1330.5) as well as electrospray mass spectrometry ($MW_{(M+H)}$=1330.6, MW_{obs}=1330.6).

Colorimetric *LacZ* assay in yeast

Cells with a freshly integrated *GAL1*::*LacZ* construct were used to assess the effects of isolated CP constructs on expression from the *GAL1* promoter[32]. Samples were prepared by growing cells overnight in raffinose media, then diluting to an OD_{600} of 0.1 and growing for 4-8 hours in galactose media to induce *LacZ* expression. Alternatively, cells were plated onto galactose plates and individual colonies were picked after 36-48 hours. Both methods produced similar results; data presented in Figure 2C are from picked colonies due to the ease with which many independent clones can be analyzed. Cells were

spun down, washed with water, and normalized to an OD_{600} of 0.4-0.8 in 110 µL in wells of a 96-well plate. 90 µL of CPRG solution (100 mM HEPES, pH 7.25, 150 mM NaCl, 0.65 mg/mL L-aspartate, 0.01 mg/mL BSA, 0.05% Tween, 0.5% SDS, 0.75 mg/mL chlorophenol red-beta-D-galactopyranoside (Sigma)) was added to each well. Plates were shaken at room temperature for 30-120 minutes. Plates were read at 578 nm and normalized for absorbance due to yeast optical density.

C. elegans strain generation and analysis of DA neuron degeneration

Nematodes were maintained following standard procedures[49]. The transgenic strains, UA109 {baInl1[P_{dat-1}:: a-syn; P_{dat-1}::gfp]; baEx84[P_{dat-1}::mCP_CP1; *rol-6* (su1006)]}, UA110 {baInl1[P_{dat-1}:: a-syn; P_{dat-1}::gfp]; baEx85[P_{dat-1}::mCP_CP1*; *rol-6* (su1006)]}, UA111 {baInl1[P_{dat-1}:: a-syn; P_{dat-1}::gfp]; baEx86[P_{dat-1}::mCP_CP2; *rol-6* (su1006)]}, UA112 {baInl1[P_{dat-1}:: a-syn; P_{dat-1}::gfp]; baEx87[P_{dat-1}::mCP_CP2*; *rol-6* (su1006)]}, UA119 {vtIs1[P_{dat-1}::gfp; *rol-6* (su1006)]; baEx93[P_{dat-1}::mCP_CP1; P_{unc-54}::mCherry]}, UA120 {vtIs1[P_{dat-1}::gfp; *rol-6* (su1006)]; baEx94[P_{dat-1}::mCP_CP1*; P_{unc-54}::mCherry]}, UA121 {vtIs1[P_{dat-1}::gfp; *rol-6* (su1006)]; baEx95[P_{dat-1}::mCP_CP2; P_{unc-54}::mCherry]}, and UA122 {vtIs1[P_{dat-1}::gfp; *rol-6* (su1006)]; baEx96[P_{dat-1}::mCP_CP2*; P_{unc-54}::mCherry]} were generated by directly microinjecting 50 ug/ml expression plasmids encoding cyclic peptide and 50 ug/ml *rol-6* or P_{unc-54}::mCherry marker into either the integrated UA44 {baInl1[P_{dat-1}:: a-syn; P_{dat-1}::gfp]} strain or the integrated BY200 {vtls1[P_{dat-1}::gfp; *rol-6* (su1006)]} as a control for toxicity[50]. Three independent transgenic lines of worms were generated for each strain. Age-synchronized worms were obtained and analyzed at the indicated time as described previously[20,38]. The six anterior

DA neurons (4 CEP and 2 ADE neurons) of 30 animals per trial were examined for neurodegeneration when the animals were 7 days old. After three trials, 90 animals for each line were analyzed, making a total of 270 worms analyzed for each transgenic strain. If a worm displayed at least one degenerative change (dendrite or axon loss, cell body loss), the animal was scored as exhibiting degenerating neurons[37,38]. For each trial, 30 worms were transferred onto a 2% agarose pad, immobilized with 2 mM levamisole, and analyzed using Nikon Eclipse E800 epifluorescence microscope equipped with Endow GFP HYQ filter cube (Chroma Technology, Rockingham, VT). Images were captured with a Cool Snap CCD camera (Photometrics, Tucson, AZ) driven by MetaMorph software (Universal Imaging, West Chester, PA).

REFERENCES

1. Driggers, E. M., Hale, S. P., Lee, J. & Terrett, N. K. The exploration of macrocycles for drug discovery - an underexploited structural class. *Nature Reviews Drug Discovery* **7**, 608-624 (2008).
2. Burja, A. M., Banaigs, B., Abou-Mansour, E., Burgess, J. G. & Wright, P. C. Marine cyanobacteria - a prolific source of natural products. *Tetrahedron* **57**, 9347-9377 (2001).
3. Kehoe, J. W. & Kay, B. K. Filamentous phage display in the new millennium. *Chemical Reviews* **105**, 4056-4072 (2005).
4. Litovchick, A. & Szostak, J. W. Selection of cyclic peptide aptamers to HCV IRES RNA using mRNA display. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 15293-15298 (2008).
5. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C. & Benkovic, S. J. Production of cyclic peptides and proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13638-13643 (1999).
6. Horswill, A. R., Savinov, S. N. & Benkovic, S. J. A systematic method for identifying small-molecule modulators of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15591-15596 (2004).
7. Horswill, A. R. & Benkovic, S. J. Cyclic peptides, a chemical genetics tool for biologists. *Cell Cycle* **4**, 552-555 (2005).

8. Scott, C. P., Abel-Santos, E., Jones, A. D. & Benkovic, S. J. Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chemistry & Biology* **8**, 801-815 (2001).
9. Cheng, L. et al. Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease. *Protein Science* **16**, 1535-1542 (2007).
10. Nilsson, L. O., Louassini, M. & Abel-Santos, E. Using siclopps for the discovery of novel antimicrobial peptides and their targets. *Protein and Peptide Letters* **12**, 795-799 (2005).
11. Kinsella, T. M. et al. Retrovirally delivered random cyclic peptide libraries yield inhibitors of interleukin-4 signaling in human B cells. *Journal of Biological Chemistry* **277**, 37512-37518 (2002).
12. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **319**, 916-919 (2008).
13. Mu, T. W. et al. Chemical and biological approaches synergize to ameliorate protein-folding diseases. *Cell* **134**, 769-781 (2008).
14. Estrada, L. D. & Soto, C. Inhibition of protein misfolding and aggregation by small rationally-designed peptides. *Current Pharmaceutical Design* **12**, 2557-2567 (2006).
15. Winderickx, J. et al. Protein folding diseases and neurodegeneration: Lessons learned from yeast. *Biochimica Et Biophysica Acta-Molecular Cell Research* **1783**, 1381-1395 (2008).
16. Duennwald, M. L., Jagadish, S., Muchowski, P. J. & Lindquist, S. Flanking sequences profoundly alter polyglutamine toxicity in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11045-11050 (2006).
17. Willingham, S., Outeiro, T. F., DeVit, M. J., Lindquist, S. L. & Muchowski, P. J. Yeast genes that enhance the toxicity of a mutant huntingtin fragment or alpha-synuclein. *Science* **302**, 1769-1772 (2003).
18. Outeiro, T. F. & Lindquist, S. Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* **302**, 1772-1775 (2003).
19. Cooper, A. A. et al. alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* **313**, 324-328 (2006).
20. Gitler, A. D. et al. The Parkinson's disease protein alpha-synuclein disrupts cellular Rab homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 145-150 (2008).
21. Polymeropoulos, M. H. et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047 (1997).
22. Spillantini, M. G. et al. alpha-synuclein in Lewy bodies. *Nature* **388**, 839-840 (1997).
23. Lee, V. M. Y. & Trojanowski, J. Q. Mechanisms of Parkinson's disease linked to pathological alpha-synuclein: New targets for drug discovery. *Neuron* **52**, 33-38 (2006).
24. Gitler, A. D. et al. alpha-Synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nature Genetics* **41**, 308-315 (2009).

25. Ramirez, A. et al. Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nature Genetics* **38**, 1184-1191 (2006).
26. Evans, T. C. & Xu, M. Q. Mechanistic and kinetic considerations of protein splicing. *Chemical Reviews* **102**, 4869-4883 (2002).
27. Ghosh, I., Sun, L. & Xu, M. Q. Zinc inhibition of protein trans-splicing and identification of regions essential for splicing and association of a split intein. *Journal of Biological Chemistry* **276**, 24051-24058 (2001).
28. Sun, P. et al. Crystal structures of an intein from the split dnaE gene of Synechocystis sp PCC6803 reveal the catalytic model without the penultimate histidine and the mechanism of zinc ion inhibition of protein splicing. *Journal of Molecular Biology* **353**, 1093-1105 (2005).
29. Naumann, T. A., Savinov, S. N. & Benkovic, S. J. Engineering an affinity tag for genetically encoded cyclic peptides. *Biotechnology and Bioengineering* **92**, 820-830 (2005).
30. Tavassoli, A. & Benkovic, S. J. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E-coli. *Nature Protocols* **2**, 1126-1133 (2007).
31. Bickle, M. B. T., Dusserre, E., Moncorge, O., Bottin, H. & Colas, P. Selection and characterization of large collections of peptide aptamers through optimized yeast two-hybrid procedures. *Nature Protocols* **1**, 1066-1091 (2006).
32. Yocum, R. R., Hanley, S., West, R. & Ptashne, M. Use of Lacz Fusions to Delimit Regulatory Elements of the Inducible Divergent Gal1-Gal10 Promoter in Saccharomyces-Cerevisiae. *Molecular and Cellular Biology* **4**, 1985-1998 (1984).
33. Outeiro, T. F. et al. Sirtuin 2 inhibitors rescue alpha-synuclein-mediated toxicity in models of Parkinson's disease. *Science* **317**, 516-519 (2007).
34. Ding, Y. et al. Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. *Journal of Biological Chemistry* **278**, 39133-39142 (2003).
35. Burdine, L. & Kodadek, T. Target identification in chemical genetics: The (often) missing link. *Chemistry & Biology* **11**, 593-597 (2004).
36. Tavassoli, A. & Benkovic, S. J. Genetically selected cyclic-peptide inhibitors of AICAR transformylase homodimerization. *Angewandte Chemie-International Edition* **44**, 2760-2763 (2005).
37. Cao, S. S., Gelwix, C. C., Caldwell, K. A. & Caldwell, G. A. Torsin-mediated protection from cellular stress in the dopaminergic neurons of Caenorhabditis elegans. *Journal of Neuroscience* **25**, 3801-3812 (2005).
38. Hamamichi, S. et al. Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 728-733 (2008).
39. Ellgaard, L. & Ruddock, L. W. The human protein disulphide isomerase family: substrate interactions and functional properties. *Embo Reports* **6**, 28-32 (2005).
40. Abou-Sleiman, P. M., Muqit, M. M. K. & Wood, N. W. Expanding insights of mitochondrial dysfunction in Parkinson's disease. *Nature Reviews Neuroscience* **7**, 207-219 (2006).

41. Cabrele, C., Fiori, S., Pegoraro, S. & Moroder, L. Redox-active cyclic bis(cysteinyl)peptides as catalysts for in vitro oxidative protein folding. *Chemistry & Biology* **9**, 731-740 (2002).
42. Kersteen, E. A. & Raines, R. T. Catalysis of protein folding by protein disulfide isomerase and small-molecule mimics. *Antioxidants & Redox Signaling* **5**, 413-424 (2003).
43. Olanow, C. W. in *Redox-Active Metals in Neurological Disorders* 209-223 (NEW YORK ACAD SCIENCES, New York, 2004).
44. Hoon, S. et al. An integrated platform of genomic assays reveals small-molecule bioactivities. *Nature Chemical Biology* **4**, 498-506 (2008).
45. van der Putten, H. et al. Neuropathology in mice expressing human alpha-synuclein. *Journal of Neuroscience* **20**, 6021-6029 (2000).
46. Rual, J. F. et al. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173-1178 (2005).
47. Yu, H. Y. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-110 (2008).
48. Li, S. M. et al. A map of the interactome network of the metazoan C-elegans. *Science* **303**, 540-543 (2004).
49. Brenner, S. Genetics of Caenorhabditis-Elegans. *Genetics* **77**, 71-94 (1974).
50. Nass, R., Hall, D. H., Miller, D. M. & Blakely, R. D. Neurotoxin-induced degeneration of dopamine neurons in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 3264-3269 (2002).

FIGURE LEGENDS

Figure 1. A cyclic peptide library that expresses and splices in yeast. (A) The SICLOPPS (split-intein mediated circular ligation of peptides and proteins) construct, designed by Benkovic and colleagues, encodes a single protein construct that yields a cyclic peptide (CP) after post-translational splicing. The *dnaE* intein C-terminal domain is encoded first, followed by the linker to be cyclized, the intein N-terminal domain, and finally a chitin binding domain affinity tag (not shown). Graphic was generated using the crystal structure of the post-splicing form of the *Synechocystis Sp.* PCC6803 *dnaE* intein, PBD ID 1ZD7 [28]. (B) Western blots against the affinity tag detect the presence and the

splicing state of the split intein construct. Log-phase yeast cultures expressing the control HPQ construct (HPQlog) show robust expression and roughly 50% distribution between the unspliced 26 kDa construct and the spliced 20 kDa byproduct. Yeast blotted at stationary phase (HPQstat) show complete splicing, while an HPQ variant with splice-disabling T69A/H72A mutations (HPQ*) shows no spliced byproduct in log phase. (C) Blots of log-phase cultures of yeast transformed with 20 randomly picked library members demonstrate that roughly 70% of the library encodes novel CP constructs that express and splice in yeast.

Figure 2. Identification of two CPs that selectively reduce α-syn toxicity in yeast. (A) The α-syn selection strain was transformed with the CP library, allowed to recover, and re-plated on galactose media to induce α-syn expression. Glucose and galactose plates are shown to highlight how a single robust hit clone could be isolated from >150,000 cells on a single plate. (B) Following isolation and amplification in bacteria, hit plasmids were individually transformed into the screening strain. Transformants were allowed to recover, normalized for cell number, serially diluted, and plated on galactose media. Four hits, named CP1-4, are shown along with a negative control CP plasmid (HPQ) and the positive control Mig1, a genetic repressor of the *GAL1* promoter. (C) A *GAL1-LacZ* reporter assay using the soluble β-galactosidase substrate chlorophenol red-beta-D-galactopyranoside (CPRG) demonstrates that CP1 and CP2 do not affect protein expression. Error bars show standard deviation from five independent trials. (D) Spotting assay, similar to (B), demonstrating that CP1 and CP2 do not affect toxicity in a yeast strain expressing multiple copies of Htt-103Q using the *GAL1* promoter. (E) Spotting

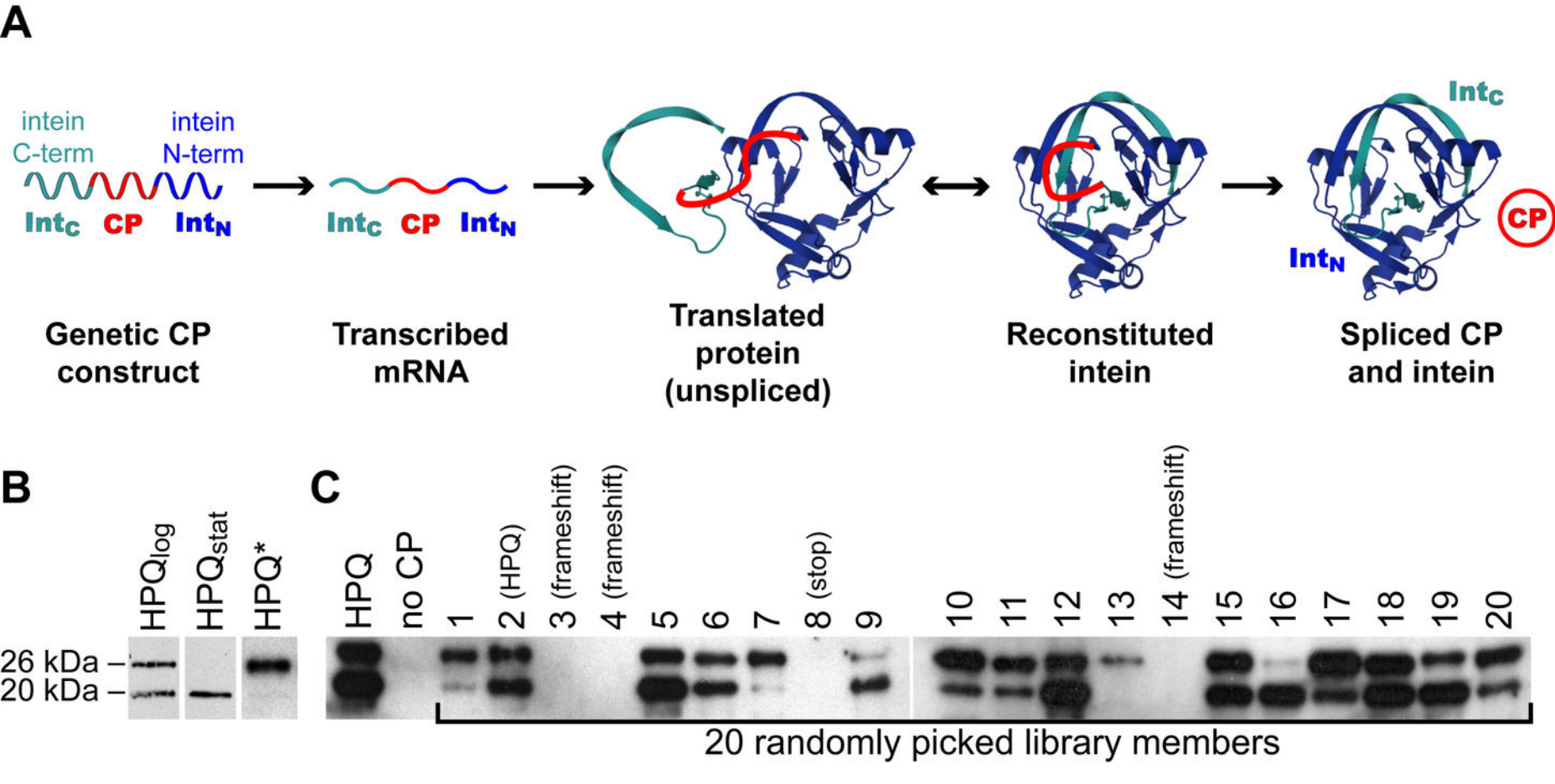
assay, similar to (B), demonstrating that splice-disabled T69A/H72A mutants of CP1-3 (denoted with asterisks) no longer suppress α-syn toxicity. Thus, for these constructs both specific CP sequences and intein processing are required for activity.

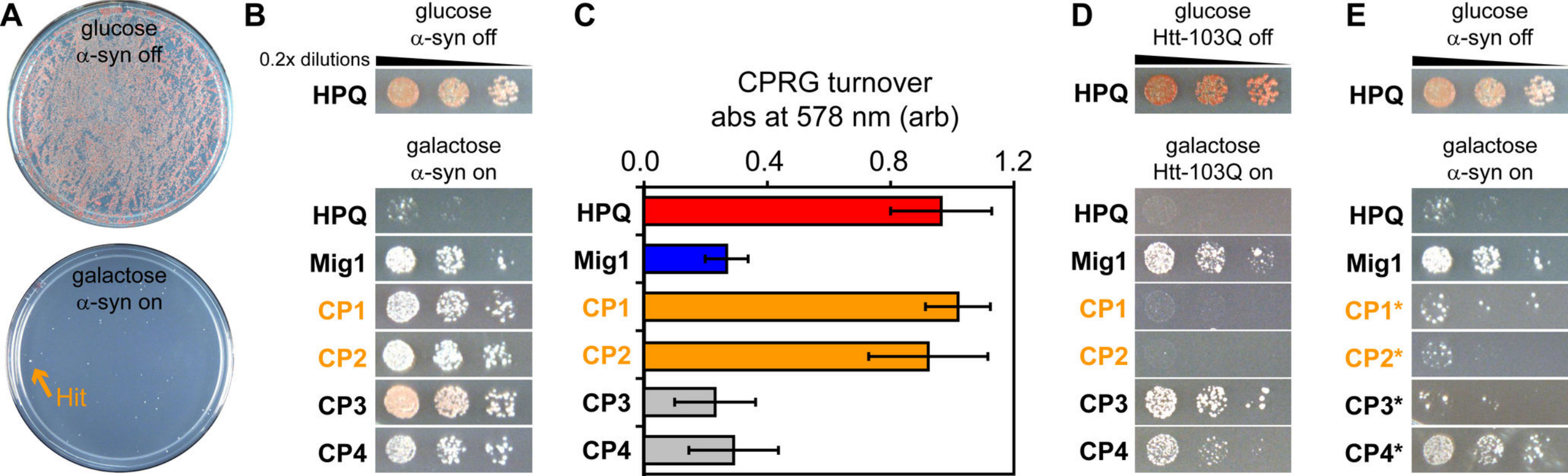
Figure 3. Rapid generation of structure-activity relationship (SAR) data and rapid minimization using point mutagenesis. (A) Sequences of the CPs encoded by CP1 and CP2. Residues found by SAR analysis to be required for function are shown in red, and residues with minor contributions are shown in pink. (B) Spotting assays of normalized, serially diluted cells of the screening strain expressing the negative control HPQ, the selected construct CP1, indicated point mutants of CP1, and the splice-disabled T69A/H72A mutant CP1*. Each construct was verified to express and splice in yeast. Similar results were obtained for CP2 and are presented in the supplement. (C) Spotting assays of yeast expressing HPQ, CP1, and additional CP1 variants. These variants illustrate incorporation of a chemical handle (here lysine substitution at position seven) as well as minimization of the encoded CP. Each construct was verified to express and splice in yeast. Asterisks denote splice-disabled T69A/H72A mutants, which were verified to express but not splice.

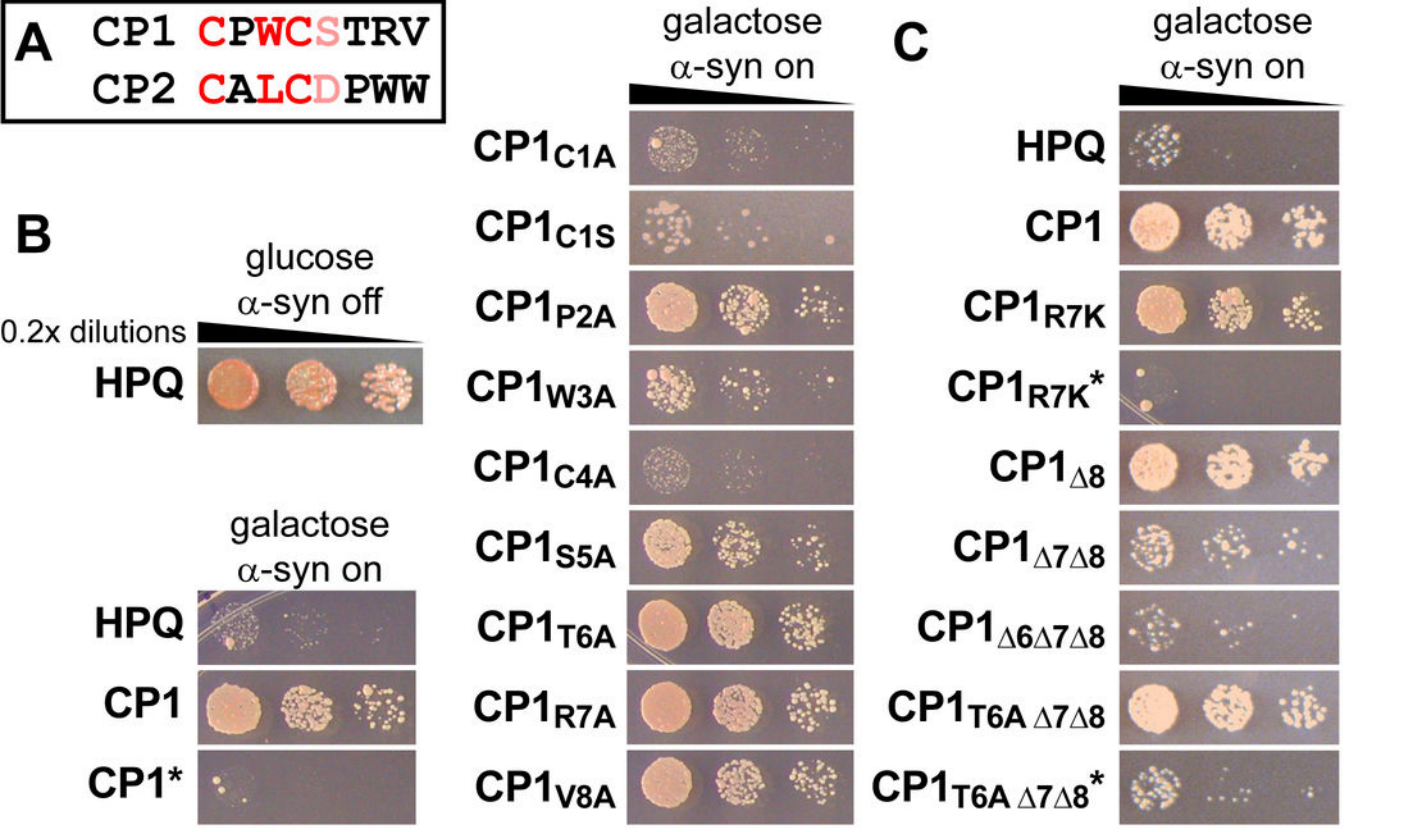
Figure 4 – Selected CPs operate downstream of known vesicle trafficking defects. (A) Representative fluorescence micrographs of cells expressing α-syn-YFP after 4 hours (first row) or 48 hours (second row), along with the indicated CP or yeast gene. (B)

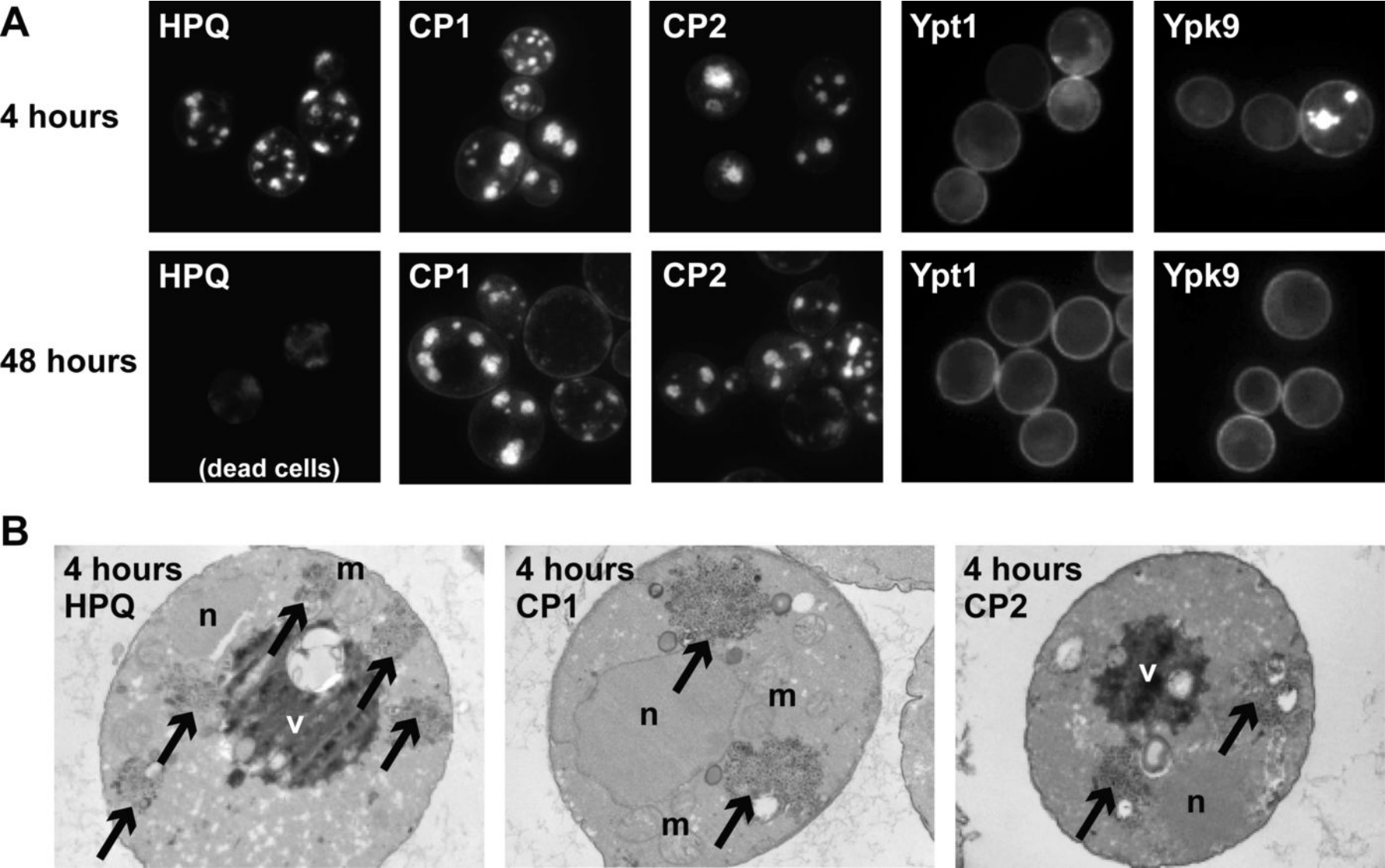
Representative electron micrographs of cells expressing α-syn-YFP after 4 hours, along with the indicated CP. m = mitochondrion, n = nucleus, v = vacuole, and arrows highlight pools of stalled vesicles. White ovals are lipid droplets.

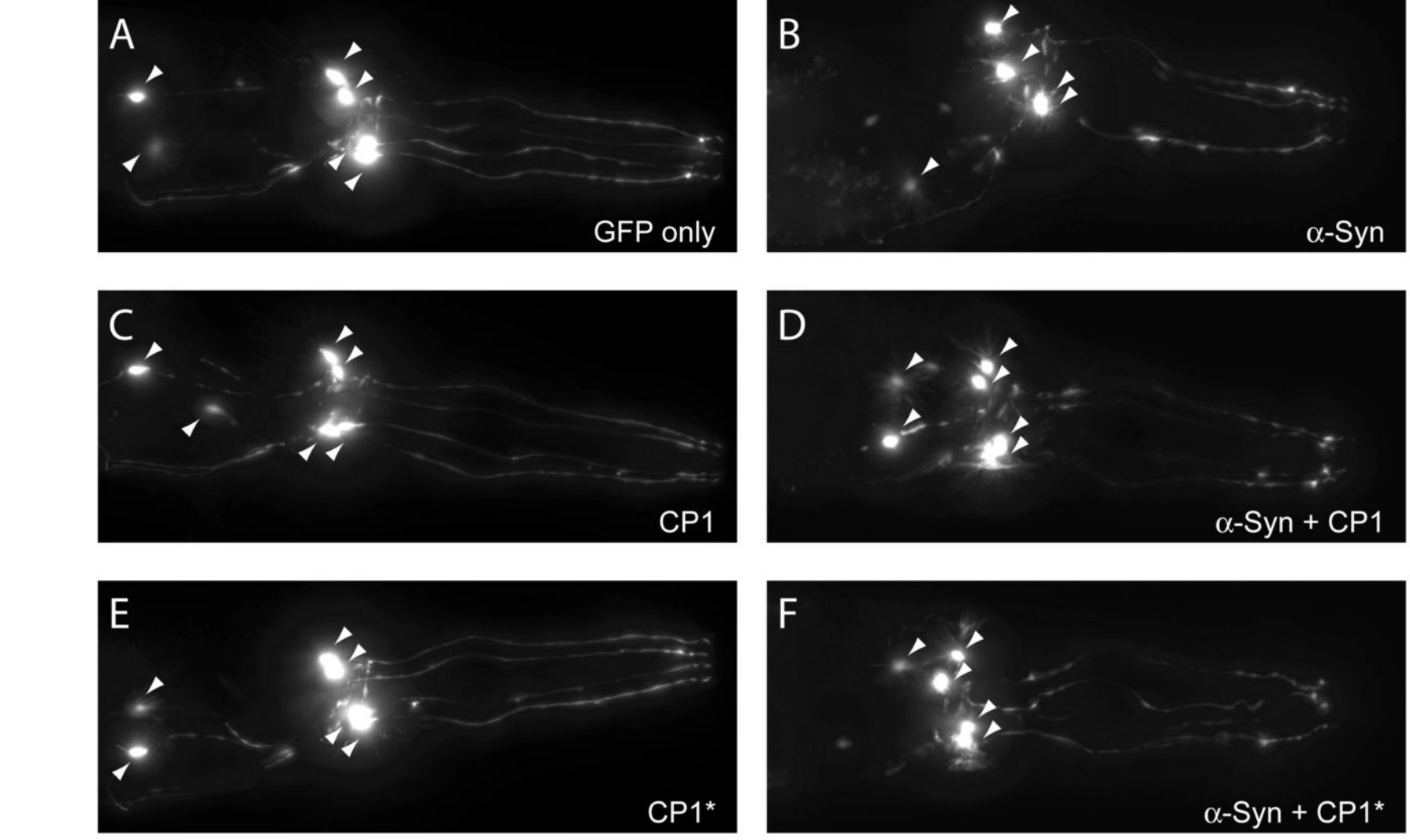
Figure 5 – Selected CPs reduce α-syn-mediated toxicity in *C. elegans* DA neurons. (A-F) Representative fluorescence micrographs of isogenic worm strains with GFP expression in DA neurons. Cell bodies are indicated by arrowheads, with processes visible to the right. (A,B) At the 7 day-old stage, worms exhibit 6 intact anterior DA neurons. However, most worms expressing α-syn have missing or degenerated anterior DA neurons. (C,D) Expression of CP1 in anterior DA neurons had no discernable effect on neuron health or development in the absence of α-syn. CP1 protected the DA neurons from α-syn-induced neurodegeneration, resulting in a greater proportion of worms with no degeneration in all 4 CEP and 2 ADE neurons. (E,F) Expression of the splice-disabled CP1* construct also had no effect on DA neurons in the absence of α-syn, but CP1* does not afford the same protection as CP1. (G) Population analysis revealed that 26% and 24% of worms expressing CP1 and CP2 showed no neurodegeneration, respectively, compared with only ~15% of control worms expressing only a-syn or expressing splice-disabled constructs CP1* and CP2* (**indicates $p<0.05$, student's t-test). Error bars show standard deviations of three trials with three independently generated worm lines, n=90 for each line, making a total of 270 animals examined for each transgenic strain. (H) Distributions of worms by number of intact, wild-type DA neurons reveals that selected CPs also reduce the severity of the observed neurodegeneration, preferentially rescuing the two more sensitive ADE neurons.

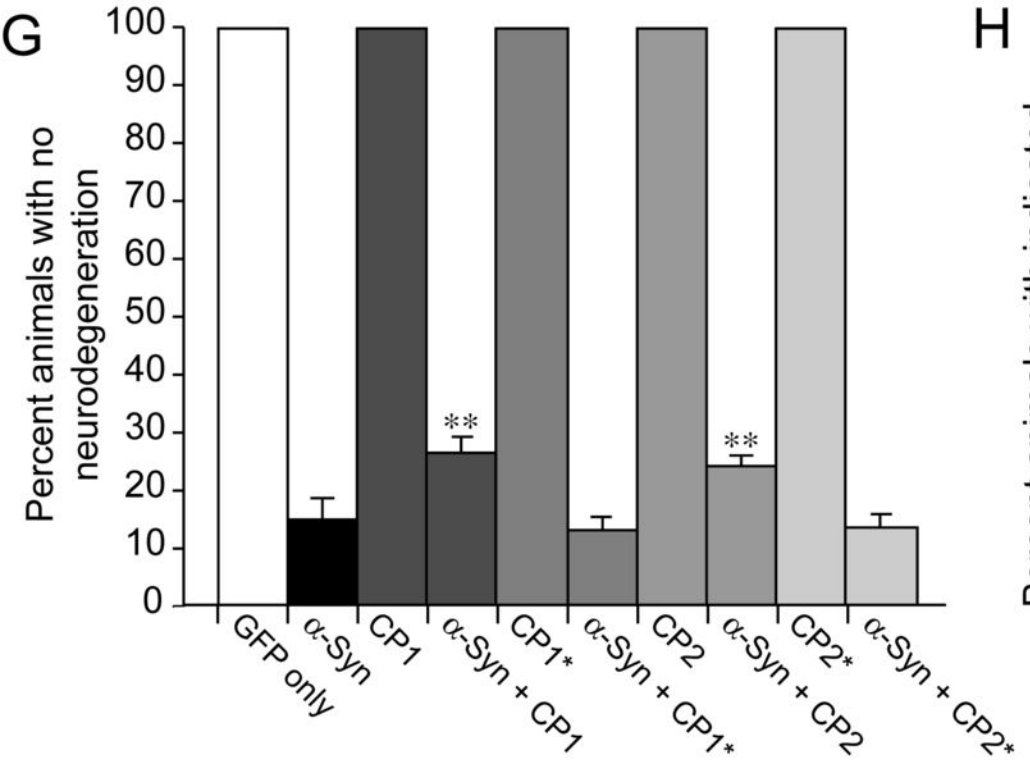

A
intein
C-term
intein
N-term
Int_C CP Int_N
Int_C CP Int_N
Int_C
CP
Int_N
Genetic CP construct
Transcribed mRNA
Translated protein (unspliced)
Reconstituted intein
Spliced CP and intein
B
HPQlog
HPQstat
HPQ*
26 kDa –
20 kDa –
C
HPQ
no CP
1
2 (HPQ)
3 (frameshift)
4 (frameshift)
5
6
7
8 (stop)
9
10
11
12
13
14 (frameshift)
15
16
17
18
19
20
20 randomly picked library members

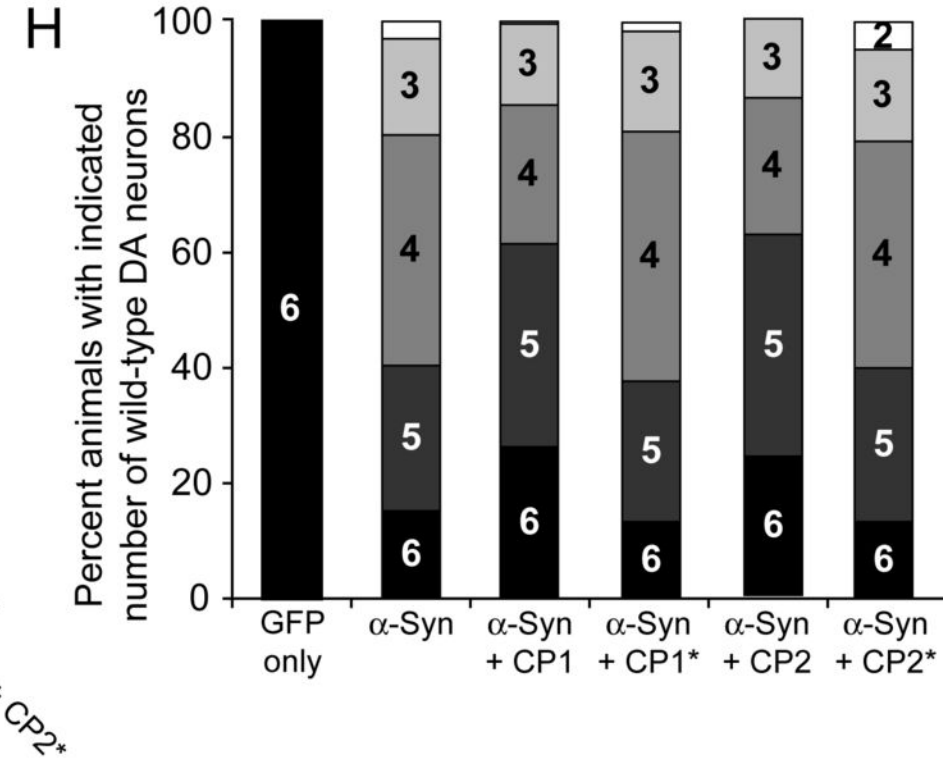

A
glucose
α-syn off
galactose
α-syn on
Hit
B
glucose
α-syn off
0.2x dilutions
HPQ
galactose
α-syn on
HPQ
Mig1
CP1
CP2
CP3
CP4
C
CPRG turnover
abs at 578 nm (arb)
0.0
0.4
0.8
1.2
HPQ
Mig1
CP1
CP2
CP3
CP4
D
glucose
Htt-103Q off
HPQ
galactose
Htt-103Q on
HPQ
Mig1
CP1
CP2
CP3
CP4
E
glucose
α-syn off
HPQ
galactose
α-syn on
HPQ
Mig1
CP1*
CP2*
CP3*
CP4*


A
CP1 CPWCSTRV
CP2 CALCDPWW
B
glucose
α-syn off
0.2x dilutions
HPQ
galactose
α-syn on
HPQ
CP1
CP1*
galactose
α-syn on
CP1C1A
CP1C1S
CP1P2A
CP1W3A
CP1C4A
CP1S5A
CP1T6A
CP1R7A
CP1V8A
C
galactose
α-syn on
HPQ
CP1
CP1R7K
CP1R7K*
CP1Δ8
CP1Δ7Δ8
CP1Δ6Δ7Δ8
CP1T6A Δ7Δ8
CP1T6A Δ7Δ8*


A
4 hours
HPQ
CP1
CP2
Ypt1
Ypk9
48 hours
HPQ
(dead cells)
CP1
CP2
Ypt1
Ypk9
B
4 hours
HPQ
m
n
v
4 hours
CP1
n
m
m
4 hours
CP2
v
n


A
GFP only
B
α-Syn
C
CP1
D
α-Syn + CP1
E
CP1*
F
α-Syn + CP1*


G
Percent animals with no neurodegeneration
100
90
80
70
60
50
40
30
20
10
0
**
**
GFP only
α-Syn
CP1
α-Syn + CP1
CP1*
α-Syn + CP1*
CP2
α-Syn + CP2
CP2*
α-Syn + CP2*


H
Percent animals with indicated number of wild-type DA neurons
100
80
60
40
20
0
6
3
4
5
6
3
4
5
6
3
4
5
6
3
4
5
6
2
3
4
5
6
GFP only
α-Syn
α-Syn + CP1
α-Syn + CP1*
α-Syn + CP2
α-Syn + CP2*