
# Mapping Signaling Mechanisms in Neurotoxic Injury from Sparsely Sampled Data Using a Constraint Satisfaction Framework

Jeffery Page[1], Kimberly A. Kelly[2], Lindsay T. Michalovicz[2], James P. O'Callahghan[2], Shichen Shen[3], Xiaoyu Zhu[3], Jun Qu[3], Jonathan Boyd[4], and Gordon Broderick[1(✉)]

[1] Center for Clinical Systems Biology, Rochester General Hospital, Rochester, NY, USA
gordonbroderick55@gmail.com
[2] Centers for Disease Control and Prevention, NIOSH, Morgantown, WV, USA
[3] Department of Pharmaceutical Sciences, University at Buffalo, Buffalo, NY, USA
[4] Department of Orthopaedic Surgery, Virginia Commonwealth University School of Medicine, Richmond, VA, USA

**Abstract.** Gulf War Illness (GWI) is a poorly understood exposure-induced neuroinflammatory disorder where complexity and the high cost of animal exposure studies has led to fragmented and sparse data sets incompatible with conventional data mining. We propose a numerical approach for generating hypotheses from sparse data to describe dysregulation of phosphoproteomic signaling in GWI brain. In an established animal model, hippocampus, and prefrontal cortex (PFC) samples were collected in mice exposed to corticosterone (CORT) to mimic high physiological stress, sarin surrogate diisopropyl fluorophosphate (DFP), CORT and DFP (CORT + DFP), as well as controls. IonStar liquid chromatography/mass spectrometry (LC/MS) profiling produced a network of 93 undirected interactions (Pearson correlation Bonferroni $< 1\%$) linking 12 hippocampal and 5 PFC phosphoproteins. With only one pre-treatment resting state and one post-treatment transient observation, conventional rate models were infeasible. Instead, a simple discrete state transition logic was applied to each network node requiring baseline be a steady state from which the network could evolve through the transient 6-h post-treatment state. Solving this as a Constraint Satisfaction (SAT) problem produced 3 competing network models where DFP directly targeted phosphorylated subspecies of sodium channel protein type 1 subunit alpha (Scn1a), protein kinase C gamma (Prkcg), sacsin molecular chaperone (Sacs), in PFC and R3H domain containing 2 (R3hdm2) in hippocampus potentiated by corticosteroids. In simulation-based searches for intervention targets inhibition of Prkcg was disproportionately represented in rescuing the model-predicted persistent illness state, though companion targets were also necessary. Results such as these suggest that a dynamically constrained model-informed design can be highly useful in the initial phases of investigation into complex poorly understood illness where detailed data is largely unavailable.

**Keywords:** Neurotoxic insult · animal model · network regulatory dynamics


D. D. Schmorrow and C. M. Fidopiastis (Eds.): HCII 2024, LNAI 14694, pp. 95–110, 2024.
https://doi.org/10.1007/978-3-031-61569-6_7

## 1 Introduction

Gulf War Illness (GWI) is a complex neuroinflammatory disorder affecting 1 of 3 Veterans of the 1990–91 Gulf War. Among the symptoms of GWI are those associated with sickness behavior (e.g. headache, fatigue, gastrointestinal distress, and neuropathic pain), suggestive of underlying neuroinflammation potentially from low-level neurotoxic exposure potentiated by in-theatre stress [1–3]. In developing a rodent model of GWI, we have shown that the exposure of mice to the stress hormone, corticosterone (CORT), at levels associated with high physiological stress, in combination with diisopropyl fluorophosphate (DFP), as a nerve agent mimic, results in marked neuroinflammation [3]. Moreover, this initial exposure to DFP can induce a "priming" of the neuroinflammatory response that persists for months in a mouse (equivalent to years in a human), consistent with the protracted sickness behavior phenotype exhibited by ill Veterans over the 26 years since their returning from theater [4–6]. Despite these important observations, the underlying molecular mechanisms of illness remain poorly understood. The complexity of the persistent GWI pathology itself coupled with the high cost of comprehensive proteomic profiling in brain tissue of exposed animals has led to fragmented and sparse data sets that are generally incompatible with conventional data mining.

The initial biological response to diverse chemical/biochemical insults is primarily coordinated by cellular signal transduction networks, which typically follow a fundamental framework: the phosphorylation/dephosphorylation cycle mediated by protein kinases and phosphatases [7–9]. Our prior work has shown that increased phosphorylation of signal transducer and activator of transcription 3 (STAT3), which activates this transcription factor, serves as a key downstream effector of proinflammatory mediators from various neurotoxic exposures [10] including CORT and DFP treatment in our GWI mouse model [1]. Indeed, exposure to these insults perturb a much broader dynamic and highly integrated network of signaling pathways responsible for normal biological function and maintenance, inducing inflammation that can progress to a chronic condition as a result of overcompensation and dysregulation [11–13]. Accordingly, an experimental approach that casts a much wider net on key phosphoproteins that mediate signal transduction through various stress and inflammatory pathways is desperately needed to elucidate the regulatory response mechanisms that support not only the onset but most importantly the dynamically stable persistence of GWI symptomatology long after insult. Only once we have identified these regulatory drivers will it be possible to identify targets that may be amenable to modulation by already available kinase/substrate inhibitors and modulators [14]. Unfortunately, mapping regulatory feedback dynamics typically requires well-sampled time course experiments making them more accessible to in vitro experimentation due to the excessive expense often associated with repeated in vivo experimentation.

Here, we demonstrate a constraint satisfaction framework where we leverage sparse experimental observations by casting individual phosphoproteins in the context of a broader regulatory network and where we apply specific hypotheses regarding the system's dynamic behavior that we expect to be true in the vicinity of these observations. Applying these additional constraints of regulatory stability to only 3 transient post-exposure observations and one pre-treatment baseline resulted in a small number of

competing models that consistently inferred phosphoproteomic profiles characteristic of stable persistent illness as well as corresponding sets of actionable targets.

## 2 Methods

### 2.1 A Small Set of Animal Experiments

All procedures were performed within protocols approved by the Centers for Disease Control and Prevention-Morgantown Institutional Animal Care and Use Committee and the US Army Medical Research and Materiel Command Animal Care and Use Review Office, and the animal facility was certified by AAALAC International. Upon arrival, mice were individually housed in a temperature- ($21 \pm 1$ °C) and humidity-controlled ($50\% \pm 10$) colony room that was maintained under filtered positive-pressure ventilation and a 12 h light (0600 EDT)/12 h dark cycle (1800 EDT) and acclimated for one week prior to the commencement of experimental procedures. Mice were given ad libitum access to food (Harlan 7913 irradiated NIH-31 modified 6% rodent chow) and sterile water. In this work, we used our acute exposure model [3] where adult male ($N = 30$, 8–12 weeks, ~ 30 g) C57BL/6 mice (Jackson Laboratory) were selectively administered the stress hormone CORT (200 mg/L in 0.6% EtOH), or not, in drinking water for 7 days then exposed to the sarin surrogate DFP by an intraperitoneal injection (4 mg/kg) or saline control (0.9%) on day 8. These mice were chosen as young adult males constituted the majority of deployed troops in the 1991 GW. Baseline control (saline, $N = 12$), CORT ($N = 6$), DFP ($N = 6$), and CORT + DFP ($N = 6$) exposed mice were sacrificed by focused wave microwave irradiation at 6 h post-exposure for all phosphoprotein analyses. Brains were removed and dissected free-hand into multiple regions, including the prefrontal cortex (PFC) and hippocampus, and immediately frozen on dry-ice and stored at −80 °C. We have chosen to examine these brain areas as they showed a significant neuroinflammatory response to these exposures in previous studies [1, 3].

### 2.2 Phosphoproteomic Profiling of Hippocampus and PFC

Samples were analyzed in triplicate using the IonStar liquid chromatography/ mass spectrometry (LC/MS) platform. Prior to analysis samples were pooled among like-treated animals to increase sample quantity and protein concentrations in each of the pooled samples measured in duplicate. The IonStar pipeline consists of two major components: i) an experimental procedure for robust sample preparation and enabling consistent, sensitive and reliable data acquisition for many biological replicates [15–17]; and ii) a data processing pipeline with optimal alignment, sensitive feature detection and stringent quality control [18–21]. IonStar supports very broad-spectrum profiling, routinely exceeding 6000 protein groups with $\geq 2$ peptide/protein in human cell/tissue samples. It also displays excellent reproducibility for low-abundance protein quantification with $< 0.2\%$ missing data, across ~ 6 orders of magnitudes in abundance. Cells were lysated by a pressurized cell in a detergent-cocktail buffer (0.5% sodium deoxycholate, 2% SDS, 2% IGEPAL® CA-630 and protease/phosphatase inhibitor cocktail), followed by a surfactant-aided/precipitation-on-pellet-digestion (SOD) method to achieve quantitative and efficient recovery of peptides, including these from hydrophobic membrane

proteins [15–17, 22]. To achieve in-depth profiling and accurate peptide ion-current quantification, digests were separated on a 100-cm-long column with 2-μm-particles by an ultra-high-pressure chromatographic setup. An Orbitrap LUMOS ultra-high-field and high-resolution mass spectrometer was then used to acquire quantitative ion-current signal and for protein identification. Individual data files were searched against the Swiss-Prot human protein database using MSGF + package leading to the quantifiable detection of 10,894 phosphorylation modifications to over 2,800 master protein species.

Phosphoprotein species of interest were those with technical replicates significantly different from saline in at least one treatment (i.e. CORT, DFP or CORT + DFP) at an adjusted t-test p-value corresponding to a Bonferroni correction $> 1\%$ and expressed with a Ln2 transformed fold change (FC) $> 2$ (Fig. 1). Calculations were performed using the Python functions *scipy.stats.ttest_ind* in the SciPy v1.11.4 library.

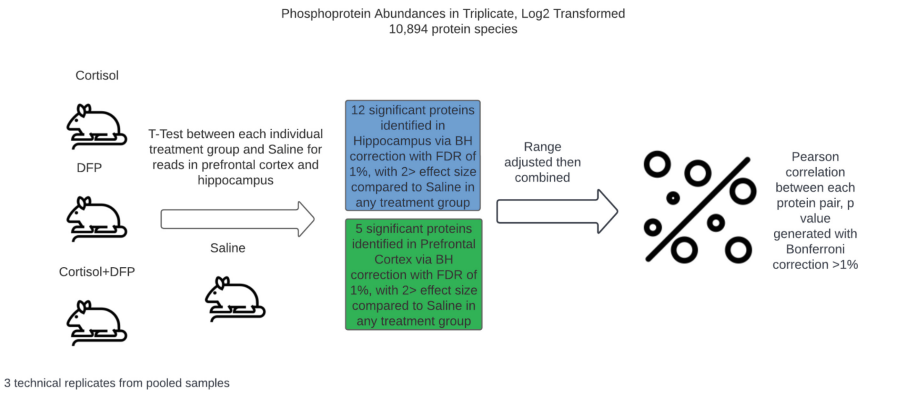


**Fig. 1.** Stringent identification of phosphoprotein species emerging in response to challenge with CORT, DFP alone and CORT + DFP in C57bl/6J male mice (Jackson Labs) at a Bonferroni Correction $<1\%$ and a signal to noise ratio $>2$. Putative interactions were inferred from simple Pearson correlation of Ln2 transformed, range-adjusted abundance values with a Bonferroni correction $>1\%$.

### 2.3 Identification of a Phosphoproteomic Regulatory Network in Brain

A first putative undirected interaction network was inferred using simple Pearson correlation as a reasonable measure of association between significantly responsive phosphoprotein species given the low number of abundance values, namely triplicate measures at each of 4 conditions. Significance of the coefficient of determination $R^2$ was based on the F statistic from the underlying linear regression computed using the *sklearn.linear_model.LinearRegression* function in the Python scikit-learn library. Raw p-values for these pairwise associations were once again corrected for multiple comparison using a Bonferroni correction with a cut off threshold of 1%. Significant undirected associations were then translated into pairs of putative opposing *directed* interactions. Of these, only those candidates directed regulatory interactions that supported a

flow of information through the network consistent with experimental observations and postulated network dynamics were retained as explained below.

Information flow through this putative directed network was modeled by applying a simple discrete decisional logic to each node originally proposed by Thomas [23] and further refined by Mendoza and Xenarios [24]. Each phosphoprotein-phosphoprotein regulatory interaction was defined by a direction, (i.e. source to target), as well as its action on a downstream target (i.e. inactivating or activating). The activation level of each phosphoprotein node was described as one of three discrete qualitative states, namely Low (0), Nominal (1), or High (2). An increase or decrease in the activation level of any given node was dictated by the states and actions of its upstream neighbors. The competing actions of upstream neighbors activated to levels above their respective perception thresholds were aggregated using a simple piece-wise linear function weighing the actions of weak inactivators against those of strong activators, and vice versa. Based on this combinatorial context-specific process an increase or decrease in the activation of each node is proposed as an update in the next iteration [25] (Fig. 2).

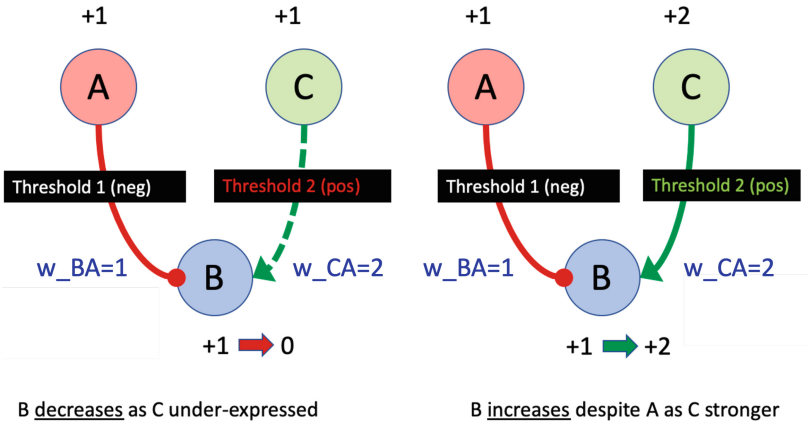


**Fig. 2.** Simple state transition logic directing the flow of information through the network. In this parameter space, each network node carries a baseline activation term and every regulatory interaction is characterized by a perception threshold and a strength of action weight.

Here, the direction (source-target) and mode of action (positive activating or negative inactivating) for each network interaction as well as logic parameters describing signal transmission thresholds and decisional weights dictating each node's state transition [26, 27] were defined as free variables and plausible values determined by solving a Constraint Satisfaction (SAT) problem [28] (Fig. 3). The expression or activation profiles observed served as both hard and soft constraints that must be met by any acceptable solution parameter set. In addition, constraints were imposed that dictate the expected dynamic behavior in the vicinity of each observation. Specifically, with respect to the observed phosphoprotein activation values at baseline (saline), we require not only that these values be recovered exactly, but also that this observed baseline state be explained by the model as a dynamically stable resting state. In other words, not only must the observed baseline state be recovered exactly, but the next predicted state must also be

identical to the current state in the absence of an external perturbation i.e. it must be explained as a steady state. These are both defined as hard constraints. In contrast, as we do not know a priori the state transition time step we establish as soft constraints the 6-hrs post-exposure phosphoprotein activation profiles to be transient observations that the most suitable network models should explain as closely as possible (Manhattan distance) in their response following each exposure. This parameter search problem was encoded using the open-source Constraint Programming and Modeling in Python (CPMpy) library in Python [29]. Solutions were generated by invoking the CP-SAT solver [30] within the Google OR-Tools optimization toolbox [31]. CP-SAT applies lazy clause generation, a hybrid approach which combines the strengths of finite domain propagation in Constraint Programming (CP) with the efficiency of Boolean Satisfiability (SAT) solvers.

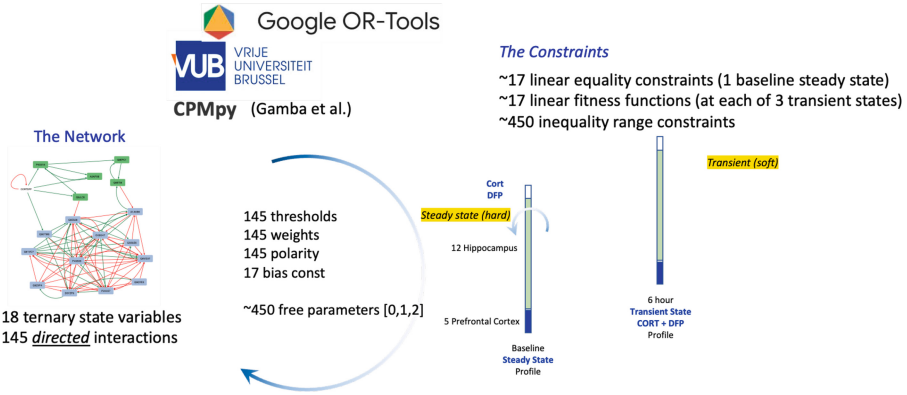


**Fig. 3.** Inferring a directed regulatory network structure and behavior from phosphoproteomic data collected in mice at baseline and at 6 h post- exposure to CORT, DFP or CORT + DFP by using this data to validate regulatory programs where the predicted dynamic responses include those observed. In essence the data is used to define a series of constraints and model parameters used as free variables in a large Constraint Satisfaction (SAT) problem.

## 2.4 Mapping Regulatory Traps and Escape Trajectories

As mentioned above, we used a discrete decisional logic to manage the signal propagation through the phosphoproteomic regulatory network and direct its evolution from one state to the next. At any given iteration, each of the network nodes is assessed for incoming signals activated above their respective perception thresholds. Based on the specific combination of active upstream mediators and their mode of action (activator or inactivator), the state of a downstream node is directed to either remain unchanged, increase, or decrease in the next iteration. Depending on the specific network update scheme, this predicted state change is applied to a single random node (asynchronous update) or to all eligible nodes simultaneously (synchronous update). Here, we used a synchronous update of all node states to improve computational efficiency and because

we are primarily interested in stable persistent behaviors. Using this framework, simulations of phosphoprotein signaling under different conditions were conducted as part of a multi-objective optimization problem directed at identifying the smallest sets of target phosphoproteins that if manipulated concurrently in a specific manner (i.e. activated or inactivated) would render the persistent illness attractor unstable and trigger a reset of phosphoprotein signaling to one ensuring a normal resting state.

Such phosphoprotein Minimal Intervention Sets (MIS) were identified by stating and solving a computationally efficient Answer Set Programming (ASP) problem [32] where idealized manipulations of network targets were iteratively assessed. As reported in previous work by our group [33], this optimization consisted of concurrently minimizing the number of target phosphoprotein nodes being manipulated, the final distance to the target state achieved by the intervention (calculated using the L1-norm), and the number of state transitions required (efficiency) to reach this treatment stabilized state. This multi-objective optimization problem was encoded in MiniZinc [34] using an in-house parser and solved with the greedy solver Chuffed [35].

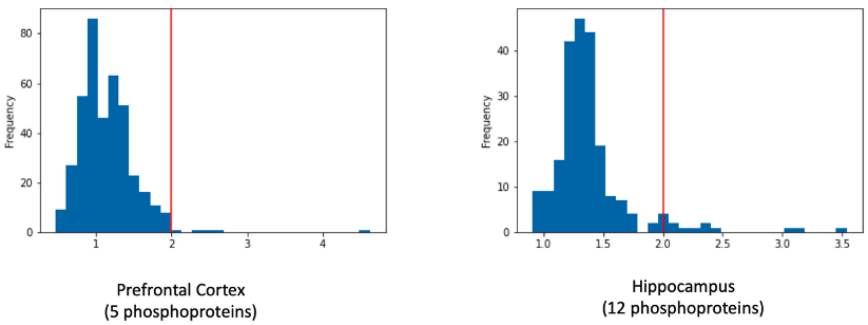


**Fig. 4.** Number of phosphoprotein species expressed significantly (Bonferroni $< 1\%$) in the hippocampus and PFC with effect size Ln2(FC) $> 2$ at 6 h pots-exposure.

## 3 Results

### 3.1 Plausible Regulatory Response Networks

We selected as nodes in the regulatory network those candidate phosphoprotein species where changes in abundance were both significant in the context of technical replication at a Bonferroni correction $< 1\%$ and important in terms of affect size with a Ln2(FC) $> 2$ (Fig. 4). We found 12 such proteins in the hippocampus and 5 in the PFC. Together with 3 insult external perturbation nodes, this resulted in a network of 20 nodes interconnected by 93 undirected associations (Pearson $R^2$ significant at Bonferroni correction $< 1\%$) (Fig. 5). Interestingly, the brain phosphoproteome targets in the PFC directly associated with the compound CORT + DFP insult consisted of the union of those targets affected by the agents CORT and DFP applied individually, suggesting a straightforward additive insult to the PFC. The 4 phosphoproteins targeted directly consisted of modifications to the master proteins Scn1, Prkcg, and Sacs in the PFC and R3hdm2 in the hippocampus.

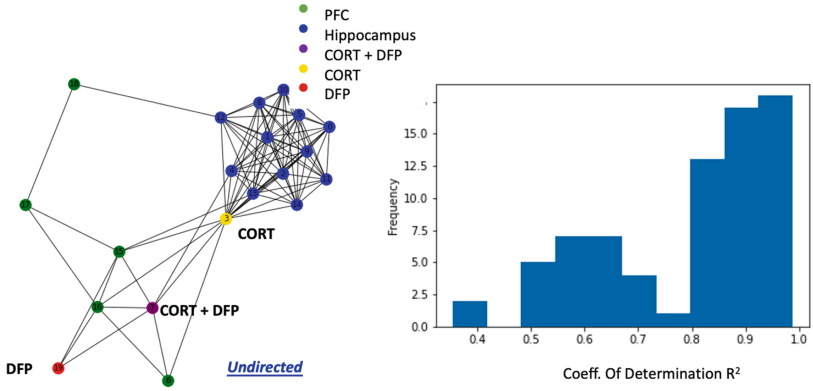


**Fig. 5.** An initial undirected association network linking 17 phosphoprotein and 3 exogenous insult nodes through 93 undirected associations.

As we are interested in the network response dynamics resulting from the propagation of the neurotoxic insult through the network, we translated all undirected associations between phosphoproteins into pairs of oppositely directed regulatory actions. The actions of exogenous agents were assumed to be unidirectional into the phosphoproteomic network. Focusing on the compounded actions of CORT + DFP, this resulted in a network of 18 nodes linked by 145 directed interactions. The corresponding parameter space consists of 17 bias terms for each of the phosphoprotein nodes and for each of the 145 directed interactions (network edges) a perception threshold, a mode of action (positive or target activation vs negative or target inactivation) as well as decisional logic weight resulting in over 450 parameters or free variables. The corresponding search for parameter values was formulated as a Constraint Satisfaction problem defined by 450 constraints on the range for each parameter, as well as 17 linear equality constraints describing the network steady state at baseline and another 17 linear inequality constraints setting an upper bound on departure from the 6-h transient response state for each of the 3 insults. Overall, the parameter search was subject to 520 constraints on 452 free variables.

Applying a lazy clause generation CP-SAT solver to this problem produced only 3 plausible solution sets where the network architecture and regulatory state transition logic could explain the baseline (saline) profile in 17 phosphoproteins as a stable steady state from which the network response to each insult evolved at 6 h to within 5% deviation (Manhattan distance) of the observed transient phosphoprotein signature (Fig. 6). In all 3 models, the direction of regulatory actions selected in the solution set supported the propagation of the neurotoxic insult through the PFC into the hippocampus. Moreover, 12 of 145 directed edges were unanimously retained across solution sets, with another 68 retained in 2 of the 3 network models (Fig. 7).

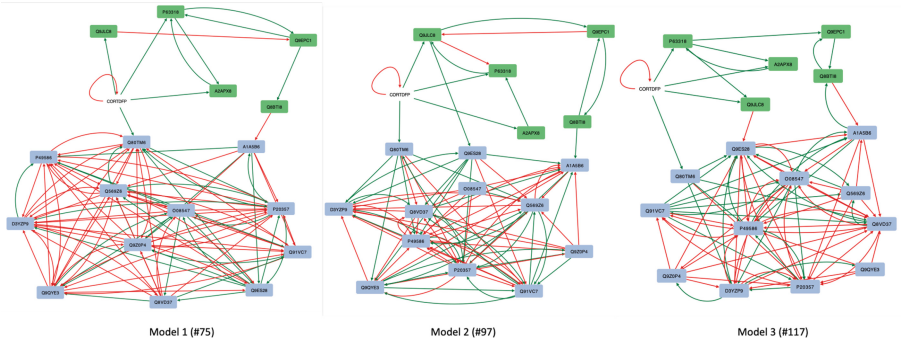


**Fig. 6.** Parameters sets identified by solving the Constraint Satisfaction (SAT) problem of Fig. 3, define 3 competing network models with $< 5\%$ deviation (Manhattan distance) from transient 6-h profile. In all 3 models information transferred to hippocampus (green nodes) from the prefrontal cortex (grey nodes), 12 of 145 directed edges unanimously retained across all 3 models, and 68 of 145 are retained in 2 of 3 solution sets (red is inactivating, green is activating). (Color figure online)

### 3.2 Predicting Persistent Illness as a Regulatory Trap

Here, we explore the idea that the persistence of chronic or very slowly evolving illness may at least in part result from an alternative homeostatic regulatory drive or coincide with a pathologic regulatory trap. To identify the regulatory traps or persistent illness conditions supported by each of the 3 candidate models above, we formulated another Constraint Satisfaction problem. This time the network structure and state transition regulatory parameters remained fixed, but we added a constraint whereby we required the regulatory network model in question to describe unknown stationary point attractors, or phosphoprotein profiles at which the next predicted network state would remain unchanged i.e. a steady state. Solving this problem for each of the 3 candidate network models identified 4 such stable point attractors or potential illness regulatory traps, one for each of models 1 and 2, and 2 for model 3. We then conducted simulations where each model was initialized with the phosphoproteomic profiles observed at 6 h following each insult to confirm which attractors would likely lie on the response path.

Of the 2 attractors supported by model 3 only one emerged as a natural endpoint following the observed responses to insult. The alternative attractor for model 3 was therefore removed from further analysis. Attractors predicted by models 1 and 2 both captured the eventual progression of insult to a persistent state (Table 1). At these candidate profiles for persistent illness, 8 of the 17 phosphoprotein species are unanimously predicted to be at their lowest activation level. Indeed, at the attractor predicted by model 2 as a persistent outcome of CORT + DFP exposure, only 2 phosphoprotein species are activated above the minimum level, namely modifications to master proteins Prkcg (P63318) and Sacs (Q9JLC8). Both are predicted to be persistently activated in the PFC at their highest level. Modifications to Srrm2 (Q8BTI8) in the PFC and Pcyt1a (P49586) in the hippocampus were also predicted to be activated in persistent states supported by models 1 and 3, offering additional potential biomarkers of chronic GWI.

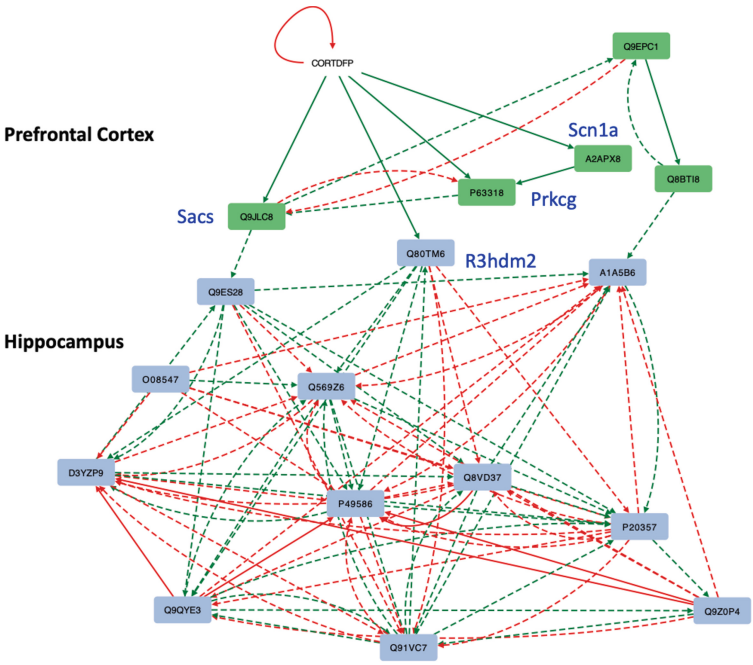


**Fig. 7.** Consensus network aggregating the 3 candidate networks showing unanimously retained directed interactions as solid lines and where red indicates inactivation and green activation of a downstream target. The combined insult of CORT and DFP directly upregulates 4 phosphoproteins, namely Scn1a, Prkcg, and Sacs in the PFC and R3hdm2 in the hippocampus. (Color figure online)

### 3.3 Predicted Rescue Strategies

In order to identify strategies for manipulating the phosphoproteome in a way that would promote an escape from the stable illness profiles described in Table 1, we formulated and solved a simulation-based Answer Set Programming (ASP) optimization problem for each model network, individually. In this approach, the model network was initialized at the corresponding predicted persistent illness attractor and sets of phosphoproteins were selected from a pool of allowable targets, and their exogenous up or downregulation were exhaustively evaluated in terms of their ability to deliver the network to the baseline normal steady state over a user-specified horizon. Here, we allowed all 17 network nodes to be selected as a target and assessed the proximity to the desired normal steady state produced by each candidate manipulation after 50 time steps. We found 59 solution sets that displaced the network state from illness to within a Manhattan distance of 20 (or $\sim$ 59% deviation) from the desired normal resting state; namely, 16 candidate manipulations to targets in model 1, 22 manipulations to targets in model 2, and 21 manipulations to targets in the phosphoproteome network of model 3 (Fig. 8). No single target solution set was identified with all solutions requiring the manipulation of at least 4 targets and only 8 solutions delivered the network back to a lasting baseline steady state (Manhattan distance of 0).

**Table 1.** Activation states in 17 network phosphoproteins at predicted long-term the persistent illness state where Low = 0, Nominal = 1, High = 2

| | Master Protein | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| **Hippocampus** | Q9QYE3 | 0 | 0 | 0 |
| | Q9ES28 | 2 | 0 | 0 |
| | O08547 | 0 | 0 | 0 |
| | P20357 | 1 | 0 | 0 |
| | Q91VC7 | 0 | 0 | 0 |
| | Q80TM6 | 1 | 0 | 0 |
| | P49586 | 1 | 0 | 2 |
| | Q8VD37 | 0 | 0 | 0 |
| | D3YZP9 | 0 | 0 | 0 |
| | Q569Z6 | 0 | 0 | 0 |
| | A1A5B6 | 1 | 0 | 0 |
| | Q9Z0P4 | 0 | 0 | 0 |
| **PFC** | P63318 | 0 | 2 | 0 |
| | Q8BTI8 | 1 | 0 | 2 |
| | A2APX8 | 0 | 0 | 2 |
| | Q9EPC1 | 0 | 0 | 0 |
| | Q9JLC8 | 0 | 2 | 0 |

Of these minimal sets, concurrent downregulation of 5, 6, 7 and 8 phosphoprotein targets in model 2 delivered the network exactly to the desired baseline self-sustaining stable resting state eliminating the need for continued treatment. The smallest set of 5 targets involved the inhibition of phosphorylated forms of Parva (Q9EPC1), Sacs (Q9JLC8), and Prkcg (P63318) in the PFC, along with inhibition of Palm (Q9Z0P4) and R3hdm2 (Q80TM6) in the hippocampus. Likewise, concurrent downregulation of 4, 5, 8 and 10 targets in model 3 exactly restored the network to the baseline steady state. The smallest of these sets consisted in the inhibition of Srrm2 (Q8BTI8) and Prkcg (P63318) in the PFC jointly with the inhibition of SGIP1 (Q8VD37) and Map2 (P20357) in the hippocampus. Importantly, all larger solution sets predicted to produce a full lasting recovery were supersets of these minimal solutions for each model, respectively.

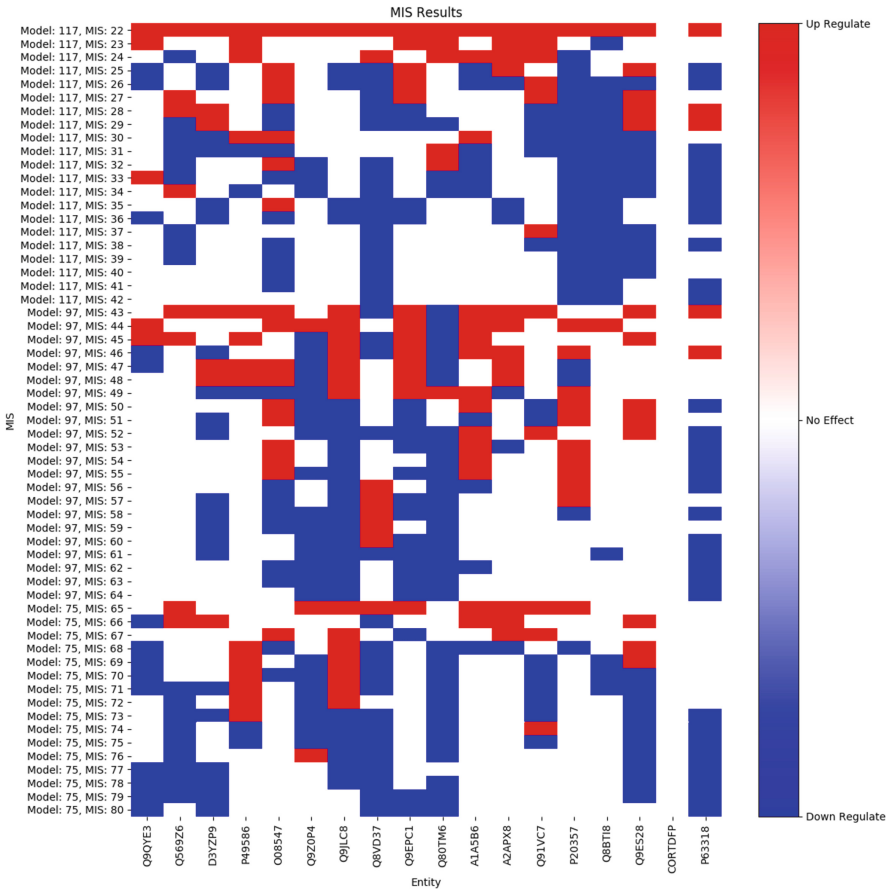


**Fig. 8.** Summary heat map of predicted escape strategies for concurrent exogenous up or down regulation of network phosphoprotein targets that would destabilize the persistent illness state (attractor) in favor of a return to stable baseline activation levels.

Looking across all intervention solutions in all 3 models, we used a simple desirability score to rank the participation of individual phosphoproteins being inhibited in those intervention sets most effective at restoring the stale baseline phosphoprotein signature (Fig. 9). This desirability score counted only those instances where a phosphoprotein target was selected for inhibition in an intervention set and weighed this instance by the squared reduction in Manhattan distance $d$ to the desired target state delivered by that intervention (i.e. [Max. $d$ – intervention residual $d$) / Max $d$]$^2$). Based on this score, inhibition of three phosphoproteins appeared especially appealing, namely Prkcg (P63318) in the PFC, as well as SGIP1 (Q8VD37) and R3hdm2 (Q80TM6) in the hippocampus. These targets were consistently selected for inhibition in those intervention sets selected to be the most successful.

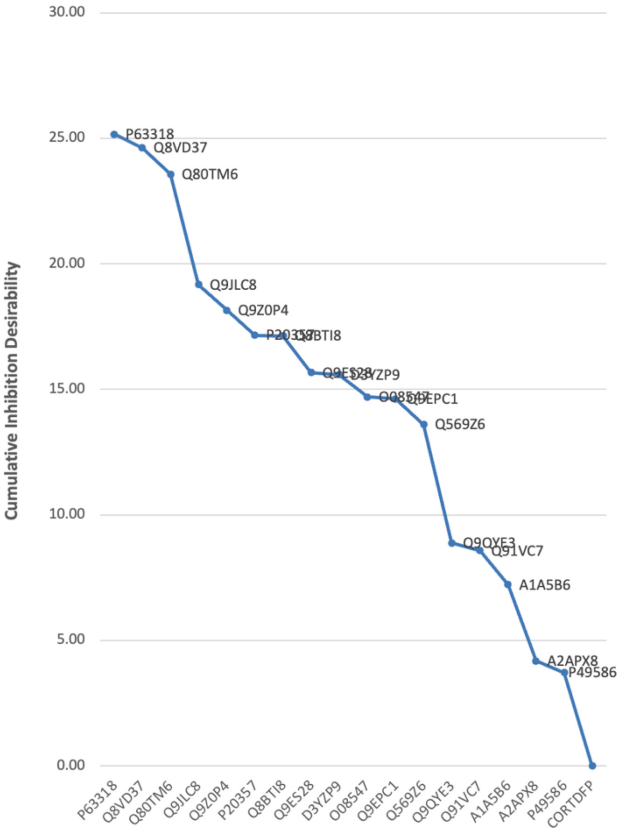


**Fig. 9.** Ranking of individual phosphoproteins selected as intervention targets based on the frequency each is selected to be inhibited in an intervention set weighted by the squared reduction in distance from the target state of baseline activation levels. Higher scores correspond to targets selected for inhibition in a greater number of more effective intervention sets.

## 4 Discussion

Here, we demonstrated a formal analytical framework whereby we augmented the use of very sparse experimental data by including and validating hypotheses regarding the expected dynamic regulatory behavior of the biological system. In the case study presented here, we have applied this approach to an early investigation of protein phosphorylation patterns in a brain induced model of Gulf War Illness. Though representing sizable resources and complex experimental procedures, the resulting data set remained small by data modeling standards. Applying this data to the validation of putative phosphoprotein regulatory network models, we identified three such models capable of explaining experimental observations and predicting plausible phosphorylation signatures for the onset and persistence of chronic illness. Moreover, using these same models and illness signatures in a simulation-based optimization scheme pointed to the inhibition of specific targets for the rescue and recovery of baseline phosphoprotein signaling. One

of the more prominent targets was protein kinase C gamma (PKCγ), Prkcg (P63318). PKC isoforms and inhibitors have gained attention in the study and treatment of several neurological illnesses, including Alzheimer's disease, ataxia, stroke, amyotrophic lateral sclerosis, and psychological disorders [36, 37].

Early results such as these are not meant to be definitive but rather offer potential avenues of study based on a systematic analysis of sparse but important new experimental data. While powerful when used appropriately, the amount and type of data needed to adequately support conventional data mining can make these methods challenging to deploy in the early stages of data collection or in small pilot studies. Here, we propose reversing the problem from one of model creation to one of model selection. By using whatever data available to validate a candidate regulatory network model rather than assemble said model de novo, the proposed approach provides an opportunity to gain early insight and perhaps even inform ongoing research studies.

**Acknowledgments.** This work was supported by Rochester Regional Health in conjunction with Elsevier BV (Amsterdam) under a collaborative research sponsorship (Broderick, PI), the US Department of Defense through Congressionally Directed Medical Research Programs (CDMRP) award GW170081 (Boyd, O'Callaghan, Kelly, Broderick, Qu) award, and Intramural funding from the Centers for Disease Control and Prevention - National Institute for Occupational Safety and Health (O'Callaghan, Kelly, Michalovicz). Pathway Studio (© 2020), Elsevier Text Mining, Elsevier Knowledge Graph and EmBio are trademarks of Elsevier Limited. Copyright Elsevier Limited except certain content provided by third parties. We wish to acknowledge the excellent technical support of Brenda K. Billig, Christopher M. Felton, and Ali A. Yilmaz.

**Mandatory Disclosure.** The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the US Department of Veterans Affairs, the US Department of Defense, Rochester Regional Health, or Elsevier BV. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention. This work was supported by the Assistant Secretary of Defense for Health Affairs through the Gulf War Illness Research Program. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Locker, A.R., Michalovicz, L.T., Kelly, K.A., Miller, J.V., Miller, D.B., O'Callaghan, J.P.: Corticosterone primes the neuroinflammatory response to Gulf War Illness-relevant organophosphates independently of acetylcholinesterase inhibition. J. Neurochem. **142**(3), 444–455 (2017)
2. Michalovicz, L.T., Kelly, K.A., Sullivan, K., O'Callaghan, J.P.: Acetylcholinesterase inhibitor exposures as an initiating factor in the development of Gulf War Illness, a chronic neuroimmune disorder in deployed veterans. Neuropharmacology **171**, 108073 (2020)

3. O'Callaghan, J.P., Kelly, K.A., Locker, A.R., Miller, D.B., Lasley, S.M.: Corticosterone primes the neuroinflammatory response to DFP in mice: potential animal model of Gulf War illness. J. Neurochem. **133**(5), 708–721 (2015)
4. Carrera Arias, F.J., et al.: Modeling neuroimmune interactions in human subjects and animal models to predict subtype-specific multidrug treatments for Gulf War illness. Int. J. Mol. Sci. **22**(16), 8546 (2021)
5. Michalovicz, L.T., Kelly, K.A., Miller, D.B., Sullivan, K., O'Callaghan, J.P.: The β-adrenergic receptor blocker and anti-inflammatory drug propranolol mitigates brain cytokine expression in a long-term model of Gulf War illness. Life Sci. **285**, 119962 (2021)
6. White, R.F., et al.: Recent research on Gulf War illness and other health problems in veterans of the 1991 Gulf War: effects of toxicant exposures during deployment. Cortex **74**, 449–475 (2016)
7. Kholodenko, B.N.: Cell-signalling dynamics in time and space. Nat. Rev. Mol. Cell Biol. **7**(3), 165–176 (2006)
8. Newman, R.H., et al.: Construction of human activity-based phosphorylation networks. Mol. Syst. Biol. **9**, 655 (2013)
9. Boyd, J.W., Neubig, R.R. (eds.): Cellular Signal Transduction in Toxicology and Pharmacology: Data Collection, Analysis, and Interpretation. John Wiley & Sons, New York (2019)
10. O'Callaghan, J.P., Kelly, K.A., VanGilder, R.L., Sofroniew, M.V., Miller, D.B.: Early activation of STAT3 regulates reactive astrogliosis induced by diverse forms of neurotoxicity. PLoS ONE **9**(7), e102003 (2014)
11. Kholodenko, B.N., Hancock, J.F., Kolch, W.: Signalling ballet in space and time. Nat. Rev. Mol. Cell Biol. **11**(6), 414–426 (2010)
12. von Kriegsheim, A., et al.: Cell fate decisions are specified by the dynamic ERK interactome. Nat. Cell Biol. **11**(12), 1458–1464 (2009)
13. Vrana, J.A., Currie, H.N., Han, A.A., Boyd, J.: Forecasting cell death dose-response from early signal transduction responses in vitro. Toxicol. Sci. **140**(2), 338–351 (2014)
14. Vrana, J.A., Boggs, N., Currie, H.N., Boyd, J.: Amelioration of an undesired action of deguelin. Toxicon **74**, 83–91 (2013)
15. Duan, X., et al.: A straightforward and highly efficient precipitation/on-pellet digestion procedure coupled with a long gradient nano-LC separation and Orbitrap mass spectrometry for label-free expression profiling of the swine heart mitochondrial proteome. J. Proteome Res. **8**(6), 2838–2850 (2009)
16. An, B., Zhang, M., Johnson, R.W., Qu, J.: Surfactant-aided precipitation/on-pellet-digestion (SOD) procedure provides robust and rapid sample preparation for reproducible, accurate and sensitive LC/MS quantification of therapeutic protein in plasma and tissues. Anal. Chem. **87**(7), 4023–4029 (2015)
17. Nouri-Nigjeh, E., et al.: Highly multiplexed and reproducible ion-current-based strategy for large-scale quantitative proteomics and the application to protein expression dynamics induced by methylprednisolone in 60 rats. Anal. Chem. **86**(16), 8149–8157 (2014)
18. Tu, C., et al.: Large-scale, ion-current-based proteomics investigation of bronchoalveolar lavage fluid in chronic obstructive pulmonary disease patients. J. Proteome Res. **13**(2), 627–639 (2014)
19. Shen, X., Hu, Q., Li, J., Wang, J., Qu, J.: Experimental null method to guide the development of technical procedures and to control false-positive discovery in quantitative proteomics. J. Proteome Res. **14**(10), 4147–4157 (2015)
20. Tu, C., et al.: Ion-current-based proteomic profiling of the retina in a rat model of Smith-Lemli-Opitz syndrome. Mol. Cell. Proteomics **12**(12), 3583–3598 (2013)

21. Tu, C., Li, J., Sheng, Q., Zhang, M., Qu, J.: Systematic assessment of survey scan and MS2-based abundance strategies for label-free quantitative proteomics using high-resolution MS data. J. Proteome Res. **13**(4), 2069–2079 (2014)
22. Shen, S., et al.: Ion-current-based temporal proteomic profiling of Influenza-a-virus-infected mouse lungs revealed underlying mechanisms of altered integrity of the lung microvascular barrier. J. Proteome Res. **15**(2), 540–553 (2016)
23. Thomas, R.: Regulatory networks seen as asynchronous automata: a logical description. J. Theor. Biol. **153**, 1–23 (1991)
24. Mendoza, L., Xenarios, I.: A method for the generation of standardized qualitative dynamical systems of regulatory networks. Theor. Biol. Med. Model. **3**(1), 1–18 (2006)
25. Sedghamiz, H., Morris, M., Craddock, T.J.A., Whitley, D., Broderick, G.: High-fidelity discrete modeling of the HPA axis: a study of regulatory plasticity in biology. BMC Syst. Biol. **12**(1), 76 (2018)
26. Sedghamiz, H., Chen, W., Rice, M., Whitley, D., Broderick G.: Selecting optimal models based on efficiency and robustness in multi-valued biological networks. In: 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 200–205. IEEE, New York (2017)
27. Sedghamiz, H., Morris, M., Craddock, T.J.A., Whitley, D., Broderick, G.: Bio-modelchecker: using bounded constraint satisfaction to seamlessly integrate observed behavior with prior knowledge of biological networks. Front. Bioeng. Biotechnol. **7**, 48 (2019)
28. Barták, R.: Constraint programming: in pursuit of the Holy Grail. Theor. Comput. Sci. **17**(12), 555–564 (1999)
29. Guns, T.: Increasing modeling language convenience with a universal n-dimensional array, CPpy as python- embedded example. In: The 18th workshop on Constraint Modelling and Reformulation (ModRef 2019). University of Connecticut, Stamford (2019)
30. Navara, M., Petrík, M.: Generators of fuzzy logical operations. In: Nguyen, H.T., Kreinovich, V. (eds.) Algebraic Techniques and Their Use in Describing and Processing Uncertainty. SCI, vol. 878, pp. 89–112. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-38565-1_8
31. Cuvelier, T., Didier, F., Furnon, V., Gay, S., Mohajeri, S., Perron, L.: OR-tools' vehicle routing solver: a generic constraint-programming solver with heuristic search for routing problems. In: 24e congrès annuel de la société française de recherche opérationnelle et d'aide à la décision (2023)
32. Guziolowski, C., et al.: Exhaustively characterizing feasible logic models of a signaling network using Answer Set Programming. Bioinformatics **29**(18), 2320–2326 (2013)
33. Sedghamiz, H., Morris, M., Whitley, D, Craddock, T.J.A., Pichichero, M., Broderick, G.: Computation of robust minimal intervention sets in multi-valued biological regulatory networks. Front. Physiol. **10**, 241 (2019)
34. Nethercote, N., Stuckey, P.J., Becket, R., Brand, S., Duck, G.J., Tack, G.: MiniZinc: Towards a standard CP modelling language. In: Bessière, C. (ed.) CP 2007. LNCS, vol. 4741, pp. 529–543. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74970-7_38
35. Chu, G., Garcia De La Banda, M., Mears, C., Stuckey, P. J.: Symmetries, almost symmetries, and lazy clause generation. Constraints **19**, 434–462 (2014)
36. Battaini, F.: Protein kinase C isoforms as therapeutic targets in nervous system disease states. Pharmacol. Res. **44**(5), 353–361 (2001)
37. Lordén, G., Newton, A.C.: Conventional protein kinase C in the brain: repurposing cancer drugs for neurodegenerative treatment? Neuronal Signaling, **5**(4), NS20210036 (2021)

Lecture Notes in Computer Science

**Lecture Notes in Artificial Intelligence 14694**

Founding Editor

Jörg Siekmann

Series Editors

The series Lecture Notes in Artificial Intelligence (LNAI) was established in 1988 as a topical subseries of LNCS devoted to artificial intelligence.

The series publishes state-of-the-art research results at a high level. As with the LNCS mother series, the mission of the series is to serve the international R & D community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings.

Dylan D. Schmorrow · Cali M. Fidopiastis
Editors

# Augmented Cognition

18th International Conference, AC 2024
Held as Part of the 26th HCI International Conference, HCII 2024
Washington, DC, USA, June 29 – July 4, 2024
Proceedings, Part I

Springer

*Editors*
Dylan D. Schmorrow
Soar Technology Inc.
Orlando, FL, USA

Cali M. Fidopiastis
Katmai Government Services
Orlando, FL, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-031-61568-9 ISBN 978-3-031-61569-6 (eBook)
https://doi.org/10.1007/978-3-031-61569-6

LNCS Sublibrary: SL7 – Artificial Intelligence



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland