STEROID: In Silico Heuristic Target Combination Identification for Disease-Related Signaling Networks

Huey Eng Chua[§]   Sourav S Bhowmick[§]   Lisa Tucker-Kellogg[†]   C. Forbes Dewey, Jr.[‡]

[§]School of Computer Engineering, Nanyang Technological University, Singapore
[†]Mechanobiology Institute, National University of Singapore, Singapore
[‡]Division of Bioengineering, Massachusetts Institute of Technology, USA

`CHUA0530|assourav@ntu.edu.sg, LisaTK@nus.edu.sg, cfdewey@mit.edu`

October 4, 2012

Abstract

Given a signaling network, the *target combination identification problem* aims to predict efficacious and safe target combinations for treatment of a disease. State-of-the-art *in silico* methods use Monte Carlo simulated annealing (MCSA) to modify a candidate solution stochastically, and use the Metropolis criterion to accept or reject the proposed modifications. However, such stochastic modifications ignore the impact of the choice of targets and their activities on the combination's *therapeutic effect* and *off-target effects* which directly affect the solution quality. In this paper, we present STEROID, a novel method that addresses this limitation by leveraging two additional heuristic criteria to minimize *off-target effects* and achieve *synergy* for candidate modification. Specifically, *off-target effects* measure the unintended response of a signaling network to the target combination and is generally associated with toxicity. *Synergy* occurs when a pair of targets exerts effects that are greater than the sum of their individual effects, and is generally a beneficial strategy for maximizing effect while minimizing toxicity. Our empirical study on the cancer-related MAPK-PI3K network demonstrates the superiority of STEROID in comparison to MCSA-based approaches. Specifically, STEROID is an order of magnitude faster and yet yields biologically relevant synergistic target combinations with significantly lower off-target effects.

1 Introduction

Despite considerable progress in genome- and proteome-based high-throughput screening methods and rational drug design, past 30 years have witnessed a steady decline in the number of new drug targets. Most modern searches for new drugs adopts the one-target, one-drug paradigm, where the focus is to identify a single new chemical entity that inhibits one well-defined molecular target. However, most diseases of interest involve physiological processes controlled in a combinatorial fashion. Specifically, redundancy and multi-functionality of biological processes are often implicated in diseases [13] such as cancer. For example, cell proliferation is used the combined control of multiple growth factor receptor pathways, and genetic experiments reveal that inhibition of any single receptor is only partially effective in blocking growth.

One way to address these challenges is through *combination therapy* by targeting multiple molecules simultaneously in a disease-related signaling network. Sometimes such strategy yields better benefits compared to a single molecule (monotherapy) for complex diseases, for dynamically changing diseases, or for diseases with a heterogeneous population of pathological mechanisms [11]. Even for diseases that are caused entirely by disruption of a single pathway (lacking in dynamics or heterogeneity), combination therapy might still offer benefits over monotherapy by virtue of spreading out the side effects to sub-toxic levels, while concentrating the desired effects on the target pathway. However, not all combination therapies produce better effects than monotherapies. For instance, in a study of combinations of analgesic drugs, some combinations (*e.g.*, `aspirin` and `pentazocine`) were beneficial, while others (*e.g.*, `acetaminophen` and `pentazocine`) were detrimental [28]. Hence, it is important to formulate strategies to develop good drug combinations which maximize the overall *therapeutic effect* while minimizing the *off-target effects*.

The identification of good drug combinations typically involves two key steps, namely identification of good target combinations and identification of appropriate set of drugs hitting these targets. In this paper, we address the *target combination identification problem*, which is complex and non-linear. Informally, this problem involves finding suitable sets of drug targets and the required *target activities* (type and extent of perturbations) for these targets for a given signaling network and a therapeutic goal. The complexity of the biological network (numerous potential drug targets and wide range of *target activities*) makes performing exhaustive search for sets of targets technically challenging, expensive and time consuming since the number of testable combinations increases exponentially with the number of variables associated with the network. Current techniques for target identification and combination therapy are generally based on empirical clinical experience. Hence, tools that can facilitate early detection of inefficacious and/or toxic target combinations *in silico* can serve as a powerful discovery and prescreening platform when coupled with other complementary technologies such as high-throughput screening.
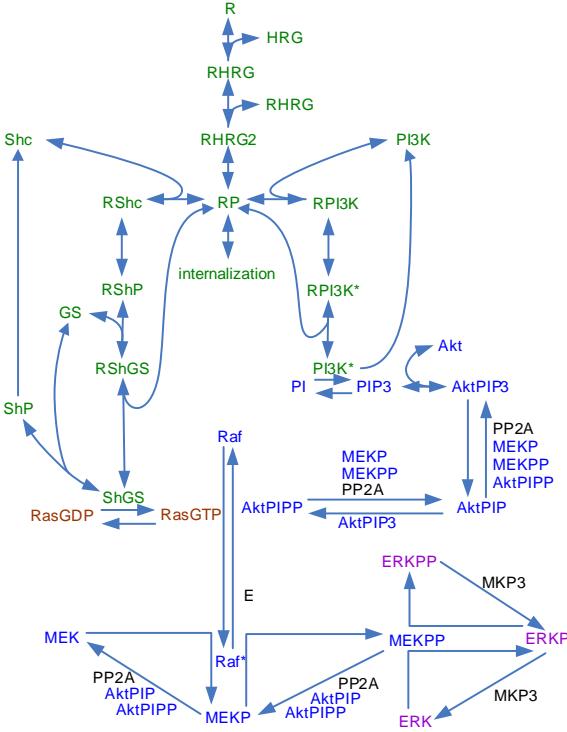
Figure 1: `MAPK-PI3K` signaling cascade [18].

Informally, the *therapeutic effect* and the *off-target effects* are measures of the intended and the unintended response, respectively, of a biological signaling network[1] to the drug combination. Since a signaling network can be modeled using mass action kinetics where ordinary differential equations (ODE) are used to model dynamics of the network, each drug effect can be simulated *in silico* by modifying appropriate signaling network model parameters. The intended response is the resulting changes to the concentration of the *output node*, while the unintended response is the resulting changes to the sum of the concentration of the rest of the nodes in the network. An *output node* is a molecule that is either involved in some dysregulated biological processes implicated in a disease, or is of interest due to its potential role in the disease. An example of an output node in the `MAPK-PI3K` network (Fig. 1) implicated in cancer is phosphorylated `ERK` (`ERKPP`) [44]. The therapeutic effect of the drug combination, `AZD6244` (`MEK` inhibitor) and `sorafenib` (`Raf` inhibitor), can be described as reducing [`ERKPP`], the node concentration we seek to decrease [12,48].

---

[1]The biological signaling processes are often modeled as hypergraphs ($G = V, E$) in systems biology, where the nodes V represent molecules (*e.g.*, proteins) and the edges E represent interactions [25]. For example, Fig. 1 depicts the hypergraph representation of the `MAPK-PI3K` signaling cascade [18].

1.1 Related Work and Motivation

Few designs of target combinations are automated. These state-of-the-art *in silico* methods are based on sequential decoding (SD) algorithms [5] or Monte Carlo simulated annealing (MCSA) [21, 47]. Stochastic search algorithms such as MCSA are expected to perform better than SD for non-linear problems [5]. MCSA modifies a candidate solution stochastically and the proposed modification is accepted or rejected using the Metropolis criterion [32]. Although stochastic candidate modification effectively covers the search space by producing a wide variety of candidates, it has two key limitations when used for identifying target combinations. First, drug targets in real signaling networks influence the therapeutic and off-target effects differently, due to one or more downstream nodes' involvement in other protein-protein interactions [9]. Ignoring this consideration may yield combinations satisfying the user-desired therapeutic effect, but with excessive off-target effects. Note that in [47] a user needs to specify *apriori* specific side effects (as input to the algorithm) in terms of the ratio of concentration of two relevant nodes. Due to the complexity of biological networks, such strategy is often impractical as it is highly unlikely for a user to know all *system-wide* side effects ahead of time. Second, although the target activity affects the combination effects, it is chosen randomly in MCSA. A judicious selection process can provide us an opportunity to improve efficiency of the overall process.

1.2 Overview and Contributions

In this paper, we present a novel and generic approach called STEROID (HeuriSTic-Based SynErgistic TaRget COmbination IDentifier) to address the aforementioned limitations. Instead of only modifying the drug target and target activity stochastically (*e.g.*, [21, 47]), STEROID judiciously modifies candidate solutions by leveraging two additional heuristic criteria for minimizing off-target effects and achieving *synergy* (detailed in Section 4). *Off-target effects* measure the unintended response of a signaling network to the drug combination and are generally associated with toxicity. *Synergy* occurs when a pair of targets exerts effects that are greater than the sum of their individual effects, and is generally a beneficial strategy for maximizing effect while minimizing toxicity. For instance, medullary thyroid cancer cells treated with `AZD6244` and `sorafenib` had better outcome in terms of cell survival and apoptosis due to drug synergy [26]. STEROID uses heuristics based on *target prioritization* methods (*e.g.*, sensitivity analysis [45, 50] and PANI [9]) which prioritize potential targets in a given disease-related network; and *Loewe additivity isobologram analysis* (LOEWE) [49] which assesses drug interaction in a combination. Specifically, target prioritization-based heuristics are used to select more effective targets to reduce off-target effects. Off-target effects is the main reason why drugs fail, and systems biology offers the hope of improving this trend by avoiding off-target effects throughout the therapy design process. LOEWE-based heuristics is used for pruning the target activity search space to reduce computa-

5

tional cost and to ensure that targets selected are synergistic. As we shall see in Section 5, the above candidate modification strategy leads to efficient identification of superior target combinations compared to MCSA-based techniques [21, 47]. Note that the goal of this work is to identify synergistic combinations of targets with reduced off-target effects and excludes the evaluation of drug compounds that bind and regulate the target molecules.

STEROID starts with a preprocessing phase that identifies the individual *target activity* necessary to achieve the therapeutic effect which will be used for selecting appropriate *target activity* of the combination in the later phase. The ODE model describing the signaling network is also modified to facilitate simulation of the target combination effect. In the second phase (simulated annealing), new candidate modifications (*drug target* and *target activity*) are proposed using heuristics. The effects of the new candidate are simulated and assessed using the Metropolis criterion [32] (Section 4.2).

The rest of the paper is organized as follows. In Sections 2 and 3, we formally introduce the terminologies and problem definition, respectively. We present the Algorithm STEROID in Section 4. Empirical analysis of our approach is discussed in Section 5. The last section concludes the paper.

2 Background

In this section, we briefly introduce *target prioritization* and *Loewe additivity isobologram analysis* which we shall be exploiting in the sequel. We begin by briefly describing the heregulin (HRG)-induced MAPK-PI3K signaling network implicated in ovarian cancer [18] (Fig. 1), which we use as a running example because its nodes are well-studied for the roles they play when targeted with relevant drugs. Details of this ordinary differential equation (ODE) model (BIOMD0000000146) are found in Biomodels.net [27]. We selected ERKPP as the output node due to its role in ovarian cancer [44]. The desired therapeutic effect was set to 50% ERKPP down-regulation but in practice would depend on the stage of the disease. Note that in practice, the therapeutic effect is dependent on the stage of the disease and is typically measured as inhibition of certain phenotypic response (*e.g.*, cell growth) which may not be linearly correlated with the inhibition of the output node concentration. In the sequel, we assume an ordinary differential equation (ODE) model of the signaling network is available.

2.1 MAPK-PI3K Network

The coupled MAPK-PI3K network (Fig. 1) is involved in up to 30% of human cancers [30] due to its roles in cell survival signaling. Here, we briefly describe the MAPK-PI3K network as depicted in Fig. 1. This model describes the heregulin (HRG)-induced ErbB receptor signaling network in Chinese hamster ovary cells. Extracellular signals such as HRG can result in dimerization of receptor tyrosine kinases on the cell surface. This causes the intracellular portions of the receptors

to be phosphorylated, which then binds to an adaptor protein known as `growth factor receptor-bound protein 2` (Grb2). This complex binds `son of sevenless` (`SOS`), thereby activating `SOS` which facilitates exchange of membrane-bound `Ras-GDP` to `Ras-GTP` [43]. The activated `Ras-GTP` in turn binds `Raf` leading to activation and phosphorylation of `Raf` (`p-Raf`) [1]. `p-Raf` then phosphorylates and activates `MEK` which in turn phosphorylates `ERK`. Phosphorylated `ERK` is translocated from the cytoplasm to the nucleus where it activates various transcription factors (e.g., `c-Myc`) [29]. Parallel to this cascade is the `PI3K-Akt` pathway, which is activated when `HRG`-stimulated receptor (`RP`) binds to and activates `PI3K`. The activated `PI3K` phosphorylates `phosphoinositol lipids` which recruit and activate `Akt`. The `MAPK` and `PI3K-AKT` cascades interacts at the level of `Raf` and `PP2A` [19].

2.2 Target Prioritization

Target prioritization methods assign prioritization rank to individual target nodes based on certain criteria (*e.g.*, the sensitivity of the output node to each target node [45]). Several target prioritization approaches such as *local sensitivity analysis* (LSA) [45], *multi-parametric sensitivity analysis* (MPSA) [50] and PANI [9] have been recently proposed.

Sensitivity analysis assigns node rank according to the sensitivity value which is the extent of the output node perturbation divided by the extent of parameter (*e.g.*, kinetic rate constant) perturbation. LSA measures sensitivity by varying a single parameter at a time [45], and global sensitivity approaches, such as MPSA, measure sensitivity by varying multiple parameters simultaneously [50].

PANI [9], in contrast, uses network information and simple empirical scores to prioritize and rank biologically relevant target molecules in signaling networks. First, it prunes the nodes based on a reachability rule to eliminate nodes that are likely to be non-regulators. Then, it ranks the resulting nodes based on the *putative target score* of each node, which is a weighted rank aggregation of a dynamic property (*profile shape similarity distance* (PSSD)) and two structural properties (*target downstream effect* (TDE) and *bridging centrality* (BC)) of the node. PSSD identifies the most relevant upstream regulators of the output node; TDE assesses the potential impact on the network when a node is perturbed; and BC identifies nodes that bridge modular subregions in a network [20]. PANI-prioritized nodes in the `MAPK-PI3K` network (*e.g.*, `AktPIP`) are found to correlate well with known ovarian cancer drug targets [9]. Hence, we reason that they are likely to form safer and more efficacious combinations.

Remark. The goal of the aforementioned target prioritization techniques is to rank individual target nodes and hence different from our goal to identify synergistic combinations of targets with reduced off-target effects.
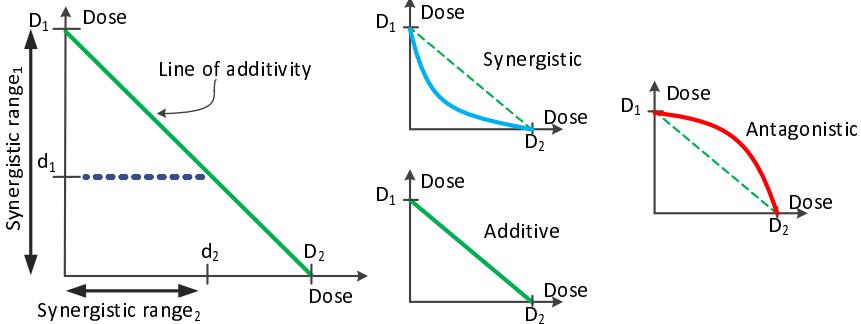
Figure 2: Isobologram. D_1 and D_2 are the dose of each drug that achieves the desired therapeutic effect if administered alone. d_1 and d_2 together can achieve the same effect.

2.3 Loewe Additivity Theory (LOEWE)

The Loewe additivity theory assumes that drugs act without self-interaction and determines drug interaction in a combination using the *combination index* [13]. Given a set of drugs X and therapeutic effect T, let D_x and d_x be the doses of drug $x \in X$ required to achieve effect T when used alone and in combination, respectively. Then, the *combination index* is defined as $CI = \sum_{x \in X} \frac{d_x}{D_x}$. The combination is *synergistic*, *additive* or *antagonistic* if $CI < 1$, $CI = 1$ or $CI > 1$, respectively. The isobologram (Fig. 2) provides a visual interpretation of LOEWE. It is a graph with the individual drug doses (D_1 and D_2) as its axes. The "line of additivity" is used to interpret the drug interaction: synergistic and antagonistic combinations are represented by drug doses that fall below and above the line of additivity, respectively [49]. We adapt this theory by replacing drugs and drug dose with *drug targets* and *target activity*, respectively for selecting synergistic *target activity* in Section 4.1.

3 Target Combination Identification Problem

In this section, we formally define the problem of *target combination identification*. We begin by introducing several concepts related to *drug target*. Henceforth, we shall use the notations given in Table 1.

3.1 Drug Target and Target Activity

A drug asserts its effect on a network through the *target*, while the *target activity* is a variable related to the extent of target perturbation. The perturbation is typically a network parameter (*e.g.*, kinetic rate constant) that controls the concentration of the node associated with the target. The drug effect is typically modeled *in silico* as modulation of the node concentration. The modulation is achieved by modifying either the node's edges (typically represented as ODE reactions) [47] or the node itself (initial concentration) depending on whether the node concentration

Symbol	Description
Γ_c	Target activity of target c.
ζ_u	Reactant-product edge set of node u.
ξ_u	Concentration-time series profile of node u.
α_u	Area under the concentration-time series profile curve of node u.
Ψ_X	Ranked list $\{\psi_{X:u_1}, \psi_{X:u_2}, \cdots, \psi_{X:u_i}\}$ based on score X where $\psi_{X:u_i}$ is the rank of node u_i.
t_c	Therapeutic effect of target c.
ρ_c	Off-target effects of target c.
λ^+	Selective pressure in the range [1–2].
δ_c	Selection probability of target c.
θ	Adjustment factor for relaxing conditions.
τ	Temperature parameter of simulated annealing.

Table 1: Notations.

varies with time. We now formally define these two concepts. We first introduce the notion of *reactant-product edge set* to facilitate exposition. Given a signaling network $G = (V, E)$ and a node $u \in V$, the *reactant-product edge set* of u is defined as $\zeta_u = R_u \bigcup P_u$ where $R_u \subset E$ and $P_u \subset E$ are the edge sets involving u as reactants and products, respectively.

Definition 1 *Given a signaling network $G = (V, E)$, and node $u \in V$ with concentration time-series profile ξ_u and reactant-product edge set ζ_u, the **drug target** of a node u is $c_{fix} = u$ if ξ_u is constant, and it is $c_{var} \in \zeta_u$ otherwise.*

Definition 2 *Given a drug target c perturbed by drug D with dissociation constant K_D, the **target activity** of c is defined as $\Gamma_c = \frac{[D]}{K_D}$ where $[D]$ is the concentration of D.*

The ODE modification varies according to the drug type (*e.g.*, activators or inhibitors) and the mechanism of action. We modeled activation using *nonessential activation* [7], and inhibition using *competitive inhibition* [47]. These reaction modifications make sense only when applied to uni-directional irreversible reactions. Note that reversible reactions can be transformed into equivalent pairs of irreversible reactions using [40].

Formally, let I be an inhibitor, A be an activator, and $c_{fix} = u$ and $c_{var} = r$ be two targets where u is a node with constant concentration time-series profile and r is a reaction in the reactant-product edge set and a node v with variable concentration time-series profile. Let $r = \frac{V_{max}[S]}{K_m + [S]}$ where V_{max} is the maximum velocity; K_m is the Michaelis-Menten constant; and $[S]$ is the concentration of the substrate S. Then, the *competitive inhibition* of c_{fix} and c_{var} are given by the

following equations:

$$\mathcal{I}(c_{fix}) = \frac{[u]_0}{\frac{[I]}{K_I}} \tag{1}$$

$$\mathcal{I}(c_{var}) = \frac{V_{max}[S]}{K_m(1 + \frac{[I]}{K_I}) + [S]} \tag{2}$$

In the above equations, $[u]_0$ is the initial concentration of u and K_I is the dissociation constant of I. Similarly, let K_A be the dissociation constant of A. The *nonessential activation* of c_{fix} and c_{var} are defined as follows.

$$\mathcal{A}(c_{fix}) = \frac{[A]}{K_A}[u]_0 \tag{3}$$

$$\mathcal{A}(c_{var}) = \frac{V_{max}[S](1 + \frac{[A]}{K_A})}{K_m + [S]} \tag{4}$$

3.2 Target Effects

Next, we formally define the notions of *therapeutic effect* and *off-target effects*. Given a signaling network $G = (V, E)$, a drug target c and the desired therapeutic effect t, let $u \in V$ be the node associated with effect t. Let α_u and α'_u be the area under the concentration-time series profile curve of node u before and after c is perturbed, respectively. Then, the *therapeutic effect* t_c and *off-target effects* ρ_c of c are given by the following equations.

$$t_c = \frac{|\alpha_u - \alpha'_u|}{\alpha_u} \tag{5}$$

$$\rho_c = \sum_{v \in V \setminus u}(\frac{|\alpha_v - \alpha'_v|}{\alpha_v}) \tag{6}$$

Note that t_c and ρ_c can be determined from *in silico* simulation using *Copasi* [40]. The combination effects are defined similarly and α can be estimated using the linear trapezoidal rule method [6].

3.3 Problem Definition

Intuitively, the goal of the *target combination identification problem* is to identify targets and their activities that achieve a user-specified therapeutic effect (*e.g.*, to achieve 50% inhibition of ERKPP) while minimizing the off-target effects. Hence, the problem can be modeled as the optimization of a *constraint satisfaction problem* (CSP) which is NP-hard [15]. The CSP is represented as a triple (X,D,C), where X, D and C represent the set of variables, the variables' domain and the set of constraints, respectively. The element X represents the set of drug targets and target activities; D represents the set of candidate targets in a given disease-related

network and the target activity range; and C represents the condition that the combination therapeutic effect must match the desired therapeutic effect. The objective of the *target combination identification problem* is to minimize the combination off-target effects.

Definition 3 *Given a set of target combination $C = \{C_1, \cdots, C_N\}$ and a desired therapeutic effect t, let $C_i = \{c_1, \cdots, c_m\}$ where $c_j \in C_i$ is the jth target in the ith combination. Let ρ_{C_i} and t_{C_i} be the off-target effects and therapeutic effect of combination C_i, respectively. Then, the **target combination identification problem** is defined as*

$$C_i = min\{\rho_{C_i} | t_{C_i} = t\}$$

4 Identifying Target Combination

In this section, we describe the heuristic algorithm STEROID for identifying target combinations. We begin by presenting the target prioritization and LOEWE-based heuristics which we shall exploit for modifying candidate solutions.

4.1 Heuristics

Target prioritization-based heuristics. The goal of using the target prioritization heuristic for target selection is to improve the average solution quality by choosing more effective targets with higher probability, thereby minimizing off-target effects. To achieve this, we first translate the node prioritization rank (Recall from Section 2.2) to an equivalent *target rank*, then convert the rank to a *selection probability* value which is used to decide if the target will be accepted. We now introduce these two concepts.

Given a signaling network $G = (V, E)$ and a target prioritization method P, let $\Psi_{P:u}$ be the rank of node $u \in V$ based on P and $\zeta_v \subset E$ be the reactant-product edge set of node $v \in V$. Let $c_{fix}, c_{var} \in C$ where C is the set of targets in G. Then, the *target ranks* of c_{fix} and c_{var} are denoted as $\Psi_{c_{fix}} = \Psi_{P:u}$ and $\Psi_{c_{var}} = \sum_{w \in W} \Psi_{P:w}$, respectively, where $W = X \bigcup Y$, $X, Y \subset V$, and $c_{var} = (X, Y)$.

The *selection probability* (δ) of a target is the likelihood of selecting the target. We use the rank-based fitness function in [3] to obtain a target's selection probability. The fitness function is based on the individual target ranks and avoids scaling problems associated with using actual objective values. The expected sampling rate of the individual target is controlled by a parameter called *selective pressure* λ^+ [46].

Observe that the aforementioned heuristic is independent of any specific target prioritization method. However, as we shall see in Section 5, PANI-based target combination identification typically generates superior quality results compared to MCSA-based techniques as the former exploits structural and dynamics properties of the signaling network [9] to improve prioritization of targets.

LOEWE-based heuristic. The effects (Section 3.2) resulting from a drug combination can be interpreted as drugs at particular dosages hitting their targets resulting in certain target activities causing a particular response of the network. Hence, an interaction of multiple targets in a combination can be assessed the same way as drug interactions (Section 2.3) by replacing the drug doses with target activities. *A target combination is guaranteed to be synergistic if the target activities are chosen from values below the line of additivity.* Following from Section 2.3, we define the *target interaction* as follows.

Definition 4 *Given a therapeutic effect t and a target combination $C = \{c_1, \cdots, c_m\}$, let $\Gamma_{0(c_i)}$ and $\Gamma_{(c_i)}$ be the target activities of the ith target in C that achieve t when targeted alone and in combination, respectively. Then, the **target combination index** of C is defined as*

$$\mathrm{TCI}_C = \sum_{c_i \in C} \frac{\Gamma_{(c_i)}}{\Gamma_{0(c_i)}}$$

*The combination is **synergistic**, **additive** or **antagonistic** if $\mathrm{TCI}_C < 1$, $\mathrm{TCI}_C = 1$ or $\mathrm{TCI}_C > 1$, respectively. The **synergistic ranges** of c_j and c_m are denoted as $[0–\Gamma_{0(c_j)})$ and $[0–\Gamma_{(c_m)})$, respectively, where $1 \leq j < m$, $\Gamma_{(c_j)} \in [0–\Gamma_{0(c_j)})$ and $\mathrm{TCI}_C < 1$.*

Graphically, the synergistic ranges of a 2-target combination can be visualized in Fig. 2 (leftmost isobologram) as "synergistic range$_1$" and "synergistic range$_2$".

4.2 The Algorithm STEROID

Systematic algorithms (*e.g.*, backtracking) that rely on partial instantiation of the candidate solutions to eliminate candidates that violate the constraints have been proposed for solving CSP [4]. In the target combination identification problem, full instantiation of the candidate solution is needed to find the combination effects. Hence, these systematic algorithms cannot be applied effectively. Meta-heuristics (e.g., SA) are used instead to find approximate solutions as they can achieve good performance performance results for large combinatorial optimization problems [34]. MCSA (a variant of SA) has been proposed for finding drug target combinations [21, 47], but suffer from certain limitations as highlighted in Section 1. In this section, we present a novel algorithm called STEROID (outlined in Algorithm 1) that addresses these limitations by leveraging on target prioritization and LOEWE-based heuristics for modifying drug target and target activity of candidate solutions.

In this section, we present STEROID (outlined in Algorithm 1) that leverages on the heuristics in Section 4.1 for modifying the candidate solutions. The inputs G and Ψ to the algorithm are used to modify the drug targets and target activities. Input G is also used to simulate the target combination effects. Note that the user can specify his preferred target prioritization method for finding Ψ. The input t is used to assess the combination effects while input $\mathcal{S}$ specifies the size

of the combination target. Several other parameters (λ^+, θ_t, θ_a, $\mathcal{N}$, τ_0 and i_{max}) that are required by STEROID are set to default values which can be modified if required (Line 2). The parameter λ^+ is used to compute the selection probability of the target. In practice, it is difficult to achieve the therapeutic effect exactly and additive target combinations are generally close to the line of additivity, but seldom "sit" exactly on it. Hence, we specify adjustment factor parameters θ_t and θ_a to relax the condition for therapeutic effect and additive combination into bound conditions, respectively (e.g., 49.5% to 50.5% inhibition of ERKPP and additive if $0.95 \leq \text{TCI} \leq 1.05$). Finally, the parameters $\mathcal{N}$, τ_0 and i_{max} are used to configure the SA and they control when the SA terminates: when $\mathcal{N}$ solutions are found or when $\tau_0 \times i_{max}$ iterations are completed.

STEROID consists of two phases, namely, the *preprocessing phase* and the *simulated annealing with heuristics phase*.

4.3 Phase 1: Preprocessing

In this phase (Line 4), the reversible reactions in G are converted into pairs of irreversible reactions using [40]. The reactions are then modified (according to Section 3.1) to simulate the effects of the targets when modulated by non-competitive inhibitors or essential activators. The set of drug targets $\mathcal{C}$ is obtained using Definition 1. The individual target activities Γ_0 required to achieve the desired therapeutic effect (50% down-regulation of ERKPP, θ_t=5%) are found using MCSA configured with the parameters τ_0 and i_{max}. Details of the individual target activity (Γ_0) are found in Table 2. Targets not listed in Table 2 cannot achieve the desired therapeutic effect alone and are deemed to have $\Gamma_0 = \infty$.

For instance, PIP3+Akt=AktPIP3 is converted into a pair of irreversible reactions using MODIFYREACTION. This pair of reactions (AktPIP3$\rightarrow$PIP3+Akt and PIP3+Akt$\rightarrow$AktPIP3) are then modified into $\dfrac{k[\text{AktPIP3}](1+\frac{[A]}{K_A})}{1+\frac{[I]}{K_I}}$ and $\dfrac{k[\text{PIP3}][\text{Akt}](1+\frac{[A]}{K_A})}{1+\frac{[I]}{K_I}}$, respectively, where k is the affinity constant; A and I are the activator and inhibitor with dissociation constants K_A and K_I, respectively. These reactions are targets ($c_{variable}$) associated with AktPIP3, Akt and PIP3 and their individual target activities are ∞ since they are not found in Table 2.

4.4 Phase 2: Simulated Annealing with Heuristics (SAH)

The SAH consists of three subphases which are repeated until either the temperature τ reaches zero or the required number of solutions $\mathcal{N}$ is found (Line 5). The subphases consist of target combination generation (the GETCOMBI procedure, Line 7); combination effects calculation (the GETEFFECT procedure, Line 8); and the test for candidate acceptance (the ACCEPTCOMBI procedure, Line 9).

In the GETCOMBI procedure (Algorithm 2), the candidate combination $\mathcal{X}$ consisting of target pairs is generated. Lines 4 to 6 implement the target prioritization-

Target	Modification	Γ_0	Target	Modification	Γ_0
Reaction 1^f	inhibition	42.975	Reaction 11	inhibition	180.193
Reaction 1^b	activation	5187.249	Reaction 12	activation	0.273
Reaction 2^f	inhibition	125.529	Reaction 13	inhibition	7.772
Reaction 2^b	activation	29.69	Reaction 14	activation	1.012
Reaction 3^f	inhibition	33.62	Reaction 15	inhibition	4.149
Reaction 3^b	activation	5857.89	Reaction 16	activation	217.788
Reaction 4	activation	47.246	Reaction 17	inhibition	2.092
Reaction 5^f	inhibition	2.081	Reaction 18	activation	1.376
Reaction 5^b	activation	10.622	Reaction 19	inhibition	1.746
Reaction 6^f	inhibition	8.957	Reaction 20	activation	13.033
Reaction 6^b	activation	16.933	Reaction 21	inhibition	2.367
Reaction 7^f	inhibition	8.144	Reaction 22	activation	2.344
Reaction 7^b	activation	8.295	Reaction 34^f	activation	3750.449
Reaction 8^f	inhibition	2.231	`MKP3`	activation	2.755
Reaction 8^b	activation	2.243	`E`	activation	2.294
Reaction 9^f	activation	0.381	`PP2A`	activation	2.315

Table 2: Details of the individual drug target achieving the desired therapeutic effect. ^{f} and ^{b} indicate forward and backward reactions, respectively.

and Line 8 LOEWE-based heuristics. The first target $\mathcal{A}$ is randomly selected using SELECTRANDOMTARGET (Line 5, Algorithm 2) and accepted in ACCEPTTARGET (Line 6) if the probability of selecting $\mathcal{A}$ (selection probability) is greater than a random number in the range [0–1] ($\delta_{\mathcal{A}} >$RAND(0,1)). Its activity is then selected within the synergistic range (Definition 4) using SELECTACTIVITY (Line 8). Similar steps are repeated to find subsequent targets and their activities.

Next, the GETEFFECT procedure (Algorithm 3) obtains the therapeutic and off-target effects by first simulating the candidate solution using Copasi (Line 1, Algorithm 3) and then calculating the therapeutic effect (Line 2, Algorithm 3) using Equation 6 and the off-target effects (Line 3, Algorithm 3) using Equation 6. These effects are used to assess the candidate in ACCEPTCOMBI (Algorithm 4) using the Metropolis criterion. A candidate is accepted under two conditions: (1) if it is synergistic, it achieves the required therapeutic effect and it has lower off-target effects lower than the current solution ($curr$) (Line 2, Algorithm 4); or (2) if it achieves the required therapeutic effect and $e^{-\frac{\rho_{\mathcal{X}}-\rho_{curr}}{\tau}} \geq$RAND(0,1) (Line 4, Algorithm 4). The UPDATESOLUTION procedure updates the solution set and the current solution with $\mathcal{X}$ if the candidate is accepted (Lines 3 and 5, Algorithm 4).

For instance, consider a two-target combination. Suppose backward reaction 29 (`AktPIP3`→`PIP3`+`Akt`) denoted as r_{29^b} is randomly selected as the first target, it will be accepted if its selection probability $\delta_{r_{29^b}} >$RAND(0,1). The activity of r_{29^b} is selected from within the range [0–$\Gamma_{0(r_{29^b})}$), where $\Gamma_{0(r_{29^b})}$ is the activity of r_{29^b} alone that is required to achieve 50% down-regulation of `ERKPP` with θ_t=5%.

Algorithm 1: Algorithm STEROID

Input: Signaling network G, prioritization rank set Ψ, therapeutic effect t, combination size $\mathcal{S}$.

Output: Solution set $\mathcal{R}$.

1   $\mathcal{R} \leftarrow$ INITIALIZE($\mathcal{R}$)
2   $(\lambda^+,\theta_t,\theta_a,\mathcal{N},\tau_0,i_{max}) \leftarrow$ SETTODEFAULTS($\lambda^+,\theta_t,\theta_a,\mathcal{N},\tau_0,i_{max}$)
3   $\tau \leftarrow \tau_0$
4   $(G,\mathcal{C},\Gamma_0) \leftarrow$ PREPROCESSINPUT($G,t,\theta_t,\tau_0,i_{max}$) /*Phase 1*/
5   **while** $\tau \geq 0$ *and* $|\mathcal{R}| \leq \mathcal{N}$ **do**
6   **foreach** *iteration i=1 to i_{max}* **do**
7          $\mathcal{X} \leftarrow$ GETCOMBI($\mathcal{C},\lambda^+,\Psi,\theta_a,\Gamma_0,\mathcal{R},\mathcal{S}$) /*Phase 2.1*/
8          $(t_\mathcal{X},\rho_\mathcal{X}) \leftarrow$ GETEFFECT($G,t,\mathcal{X}$) /*Phase 2.2*/
9          $\mathcal{R} \leftarrow$ ACCEPTCOMBI($t_\mathcal{X},\rho_\mathcal{X},t,\theta_a,\theta_t,\tau,\mathcal{R}$) /*Phase 2.3*/
10   Decrement τ

Algorithm 2: The GETCOMBI Procedure (Phase 2.1)

Input: Target candidate set $\mathcal{C}$, selective pressure λ^+, prioritization rank set Ψ, adjustment factor for target interaction θ_a, individual target activity set Γ_0, solution set $\mathcal{R}$, combination size $\mathcal{S}$.

Output: Combination candidate $\mathcal{X} = \{(x_1,\Gamma_1),\cdots,(x_S,\Gamma_S)\}$.

1   $\mathcal{X} \leftarrow$ INITIALIZE($\mathcal{X}$)
2   $\mathcal{Y} \leftarrow$ INITIALIZE($\mathcal{Y}$)
3   **foreach** *combination candidate component* $(x_i,\Gamma_i)=(x_1,\Gamma_1)$ *to* (x_S,Γ_S) **do**
4   **while** ISNULL(x_i) *is* TRUE **do**
5          $\mathcal{A} \leftarrow$ SELECTRANDOMTARGET($\mathcal{C}/\mathcal{Y},\mathcal{R}$)
6          $x_i \leftarrow$ ACCEPTTARGET($\mathcal{A},\mathcal{C}/\mathcal{Y},\Psi,\lambda^+$)
7      $\mathcal{Y} \leftarrow \mathcal{Y} \bigcup x_i$
8      $\Gamma_i \leftarrow$ SELECTACTIVITY($x_i,\mathcal{Y},\Gamma_0,\theta_a,\mathcal{R}$)

The second target (*e.g.*, r_{13} (Raf$\rightarrow$Raf*)) is randomly selected from the set of candidates excluding r_{29^b} (C/r_{29^b}) and will be accepted if $\delta_{r_{13}} >$ RAND(0,1). The activity of r_{13} is selected from the range $[0$–$\Gamma_{r_{13}})$ where $\frac{\Gamma_{(r_{29^b})}}{\Gamma_{0(r_{29^b})}} + \frac{\Gamma_{(r_{13})}}{\Gamma_{0(r_{13})}} < 1-\theta_a$ and $\Gamma_{(r_{29^b})} \in [0$–$\Gamma_{0(r_{29^b})})$. The therapeutic and off-target effects of the combination $c = ((r_{29^b}, \Gamma_{(r_{29^b})}), (r_{13}, \Gamma_{(r_{13})}))$ are computed as $t_c = \frac{|\alpha_{\text{ERKPP}} - \alpha'_{\text{ERKPP}}|}{\alpha_{\text{ERKPP}}}$ and $\rho_c = \sum_{v \in V_{\text{MAPK}} \backslash \text{ERKPP}} (\frac{|\alpha_v - \alpha'_v|}{\alpha_v})$, respectively. Finally, the Metropolis criterion is used to assess if the combination will be added into the solution set.

Theorem 1 STEROID *has time complexity $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|)$ where τ_0 is the initial temperature; i_{max} is the limit on iterations per cycle; $|E|$ is the number of irreversible reactions; and $|\xi|$ is the number of time points in the concentration time-series profiles used to estimate the target effects.*

Proof 1 In the PREPROCESSINPUT procedure (Line 4, Algorithm 1), given a net-

---

Algorithm 3: The GETEFFECT Procedure (Phase 2.2)

Input: Signaling network G, therapeutic effect t, candidate combination $\mathcal{X}$.
Output: Combination therapeutic effect $t_{\mathcal{X}}$, combination off-target effects $\rho_{\mathcal{X}}$.

1 $(\alpha, \alpha') \leftarrow$ SIMULATECOMBINATION($G,\mathcal{X}$)
2 $t_{\mathcal{X}} \leftarrow$ GETTHERAPEUTICEFFECT(α,α',t)
3 $\rho_{\mathcal{X}} \leftarrow$ GETOFFTARGETEFFECTS(α,α',t)

Algorithm 4: The ACCEPTCANDIDATE Procedure (Phase 2.3)

Input: Combination therapeutic effect $t_{\mathcal{X}}$, combination off-target effects $\rho_{\mathcal{X}}$, therapeutic effect t, adjustment factor for target interaction θ_a, adjustment factor for therapeutic effect θ_t, temperature τ, solution set $\mathcal{R}$.
Output: Solution set $\mathcal{R}$.

1 $\mathcal{M} \leftarrow$ GETCURRENTSOLUTION($\mathcal{R}$)
2 **if** $\rho_{\mathcal{X}} \leq \rho_{\mathcal{M}}$ *and* $\sum_{x \in \mathcal{X}} \frac{\Gamma_{\mathcal{R}}}{\Gamma_{\mathcal{R}0}} < 1 - \theta_a$ *and* $\frac{|t_{\mathcal{X}} - t|}{t} \leq \theta_t$ **then**
3 $\quad$ $\mathcal{R} \leftarrow$ UPDATESOLUTION($\mathcal{R},\mathcal{X}$)

4 **else if** $\frac{|t_{\mathcal{X}} - t|}{t} \leq \theta_t$ *and* RAND(0,1)$\leq e^{-\frac{\rho_{\mathcal{X}} - \rho_{\mathcal{M}}}{\tau}}$ **then**
5 $\quad$ $\mathcal{R} \leftarrow$ UPDATESOLUTION($\mathcal{R},\mathcal{X}$)

---

work model $G = (V, E)$, the process of converting the reversible reactions to irreversible pairs of reactions; of modifying the reactions to stimulate the effects of a non-competitive inhibitor or an essential activator; and of obtaining the set of drug targets all require $O(|E|)$ time. We use the MCSA to find the individual target activity in PREPROCESSINPUT. Each MCSA iteration consists of the following steps: generating random candidate solution ($O(1)$); simulating target combination effect using an ODE solver ($O(|E| \cdot |\xi|)$ [38]); estimating the therapeutic and off-target effects ($O(|\xi|)$); and determining acceptance of candidate ($O(1)$), where $|\xi|$ is the number of time points in the concentration-time series profile curve. Hence, the time complexity of each MCSA iteration is $O(|E| \cdot |\xi|)$. In the worst case, no values satisfying the desired therapeutic effect is found when the MCSA terminates on completing all runs ($\tau_0 \cdot i_{max}$ iterations) resulting in a time complexity of $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|)$ for obtaining the individual target activity for each target c. Hence, PREPROCESSINPUT has time complexity of $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|)$ since $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|) > O(|E|)$. In the SA phase, in the worst case, the SA terminates on completing all runs ($\tau_0 \cdot i_{max}$ iterations) without finding $\mathcal{N}$ solutions (Line 5, Algorithm 1).

Similar to the MCSA, each SAH iteration consists of the following steps: generating candidate solution (Line 7, Algorithm 1); simulating the combination effects using an ODE solver (Line 1, Algorithm 3); estimating the combination effects (Lines 2 to 3, Algorithm 3); and determining acceptance of the candidate solution (Line 9, Algorithm 1). Generating the candidate solution involves target prioritization-based heuristic for selecting the target. This heuristics accepts higher prioritized targets with higher probability (Section 4.1). In the worst case, the low-

PMID	Target 1	Target 2
22180401	Akt inhibitor	MEK inhibitor
21632463	PI3K inhibitor	MEK inhibitor
21062259	Akt inhibitor	MEK inhibitor
14675307	PI3K inhibitor	Akt inhibitor

Table 3: PubMed results relevant to ovarian cancer drug combinations targeting the MAPK-PI3K network.

est prioritized target (c_l) from the set of individual targets (C) is considered and will be accepted if a randomly generated number is lower than the selection probability of c_l (δ_{c_l}). According to Section 4.1, δ_{c_l} is computed using the rank-based fitness function. Hence, $\delta_{c_l} = \frac{\text{TARGETRANK}(c_l)}{\sum_{c_i \in C} \text{TARGETRANK}(c_i)}$ where $\text{TARGETRANK}(c_l)$=2-λ^++$\frac{2(\lambda^+-1)(1-1)}{(|c|-1)}$ and $\sum_{c_i \in c} \text{TARGETRANK}(c_i)$=$|C|(2-\lambda^+) + 2(\lambda^+ - 1)$. Thus, $\delta_{c_l} = \frac{(2-\lambda^+)}{|C|(2-\lambda^+)+2(\lambda^+-1)}$. Since λ^+ is in the range [1–2], $\frac{1}{|C|} \leq \delta_{c_l} \leq \frac{1}{(|C|+1)}$. This implies that it takes about $O(|C|)$ tries in order to accept the worst candidate. Hence, the complexity of generating the candidate solution is $O(|\mathcal{X}| \cdot |C|)$ where $|\mathcal{X}|$ is the size of the candidate combination and $|C|$ is the size of the target candidate. Simulating the combination effects takes $O(|E| \cdot |\xi|)$ [38]. Estimating the combination effects takes $O(|\xi|)$ and determining acceptance of the candidate solution takes $O(1)$. Since $|\xi| \gg |\mathcal{X}|$ and $O(|E|) = O(|C|)$ (Definition 1), the time complexity of SAH is $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|)$ time in the worst case. Thus, the overall time complexity of STEROID is $O(\tau_0 \cdot i_{max} \cdot |E| \cdot |\xi|)$ which is polynomial. ∎

5   Performance Study

STEROID was implemented using Java. In this section, we investigate the performance of STEROID and compare it with state-of-the-art MCSA-based techniques [21, 47]. Since STEROID involves two heuristics and it is orthogonal to any specific target prioritization technique, we use several variants of it for comparative study. We denote variants of STEROID enabled with one and two heuristics as STEROID-X and STEROID-XX, respectively, where X$\in$ {L,P,S,M} and L=LOEWE, P=PANI [9], S=LSA [45] and M=MPSA [50]. Note that modification of the candidate differs depending on the heuristics used. For MCSA, the targets and activities are selected randomly. For STEROID-P, STEROID-S and STEROID-M, the targets are selected using Algorithm 2 (Lines 4- 6) while the activities are selected randomly. For STEROID-L, the targets are selected randomly while the activities are selected from within the synergistic range. For STEROID-PL, STEROID-SL and STEROID-ML, the targets and activities are selected using Algorithm 2 (Lines 4- 8).

We ran the experiments on an Intel 2.93GHz Xeron processor machine with 12GB RAM running Microsoft Windows 7 to obtain 10 sets of results for each approach. The MAPK-PI3K network [18] was used for analysis and the desired

Target in Table 3	Inhibition Mechanism	Corresponding target(s) in **MAPK-PI3K** network
Akt inhibitor	Disruption of Akt binding to its membrane localizing factor (PIP3) [35] or dephosphorylation of PIP3 [31]	Activators of Reaction 29^b, Reaction 30, Reaction 33; Inhibitors of Reaction 29^f, Reaction 31, Reaction 32
MEK inhibitor	MEK dephosphorylation [33] or blockade of MEK phosphorylation [14]. Is also used to achieve ERK inhibition as there is no known ERK inhibitors.	Activators of Reaction 16, Reaction 18, Reaction 20 or Reaction 22; Inhibitors of Reaction 15, Reaction 17, Reaction 19 or Reaction 21
PI3K inhibitor	Inhibits PI3K in ATP-competitive manner [31]	Activators of Reaction 24^b, Reaction 26; Inhibitor of Reaction 24^f

Table 4: Mapping between targets in Table 3 and MAPK-PI3K network.

therapeutic condition was chosen as 50% ERKPP down-regulation (Section 2). For all experiments, the combination size was set to 2 and the following default values were used for the rest of the parameters: $\{\tau_0=100; i_{max}=500; \mathcal{N}=50; \theta_t=\theta_a=5\%; \lambda^+=1.8\}$. We used Copasi for estimating the combination effects and its parameters were set following [9]. In the tables presented, the terms *ACT* and *IN* denote activators and inhibitors, respectively. Forward and backward reactions are marked with superscripts ^{f} and ^{b}, respectively.

5.1 Biological Relevance

First, we examined each approach's ability to find a set of benchmark target combination relevant to ovarian cancer and targeting the MAPK-PI3K network. The benchmark combination set is curated from literature in the PubMed repository using "ovarian", "cancer", "combination" as keywords. Out of 5863 PubMed records (as of 17 March 2012), only 4 (Table 3) fitted the criteria. Table 4 shows the targets in the MAPK-PI3K network that match those in Table 3. We examined our solution set to identify those combinations (Tables 5 and 6) involving the targets in Table 4 and found that *the majority (43.6%) of the target combinations were found using* STEROID-PL. We noted that the off-target effects of these solutions were relatively low (in the range [4–26]) and were likely to be good combinations, correlating well with literature [2, 23].

Since the solutions in Tables 5 and 6 were not the best within our solution set, we went on to examine if the best solutions are biologically relevant and whether one approach is better than another for biological purposes. We pooled solutions from all approaches to identify the top-10 solutions having the least off-target ef-

No.	Target 1 [Activity]	Target 2 [Activity]	ρ	Approach	PubMed Combination
1	Reaction 33 ACT [673.356]	Reaction 18 ACT [1.282]	4.164	STEROID-L	Akt/MEK inhibitors
2	Reaction 30 ACT [8930.634]	Reaction 17 IN [1.646]	8.100	STEROID-L	Akt/MEK inhibitors
3	Reaction 30 ACT [6874.187]	Reaction 22 ACT [2.162]	7.172	STEROID-L	Akt/MEK inhibitors
4	Reaction 18 ACT [1.271]	Reaction 29^b ACT [3833.824]	6.081	STEROID-L	Akt/MEK inhibitors
5	Reaction 24^f IN [4207.992]	Reaction 16 ACT [144.239]	25.197	STEROID-L	MEK/PI3K inhibitors
6	Reaction 22 ACT [2.135]	Reaction 24^b ACT [1766.783]	7.403	STEROID-L	MEK/PI3K inhibitors
7	Reaction 30 ACT [1140.911]	Reaction 20 ACT [9.866]	8.764	STEROID-L	Akt/MEK inhibitors
8	Reaction 17 IN [1.739]	Reaction 30 ACT [383.720]	8.118	STEROID-PL	Akt/MEK inhibitors
9	Reaction 21 IN [2.175]	Reaction 24^b ACT [887.656]	5.028	STEROID-PL	MEK/PI3K inhibitors
10	Reaction 16 ACT [174.005]	Reaction 32 IN [6494.301]	11.726	STEROID-PL	Akt/MEK inhibitors
11	Reaction 17 IN [1.851]	Reaction 32 IN [574.388]	4.241	STEROID-PL	Akt/MEK inhibitors
12	Reaction 21 IN [2.241]	Reaction 29^f IN [7724.345]	5.156	STEROID-PL	Akt/MEK inhibitors
13	Reaction 16 ACT [171.804]	Reaction 33 ACT [5469.422]	11.695	STEROID-PL	Akt/MEK inhibitors
14	Reaction 33 ACT [4200.96]	Reaction 18 ACT [1.298]	4.183	STEROID-PL	Akt/MEK inhibitors
15	Reaction 24^b ACT [2982.964]	Reaction 16 ACT [193.024]	18.431	STEROID-PL	MEK/PI3K inhibitors
16	Reaction 17 IN [1.721]	Reaction 26 ACT [1375.482]	10.624	STEROID-PL	MEK/PI3K inhibitors
17	Reaction 30 ACT [3145.694]	Reaction 18 ACT [1.017]	8.112	STEROID-PL	Akt/MEK inhibitors
18	Reaction 18 ACT [1.081]	Reaction 30 ACT [1220.678]	8.177	STEROID-PL	Akt/MEK inhibitors
19	Reaction 19 IN [1.483]	Reaction 30 ACT [9385.883]	7.834	STEROID-PL	Akt/MEK inhibitors
20	Reaction 22 ACT [2.105]	Reaction 30 ACT [4633.781]	7.155	STEROID-PL	Akt/MEK inhibitors
21	Reaction 17 IN [1.810]	Reaction 33 ACT [496.541]	4.209	STEROID-PL	Akt/MEK inhibitors
22	Reaction 17 IN [1.722]	Reaction 26 ACT [6711.221]	10.638	STEROID-PL	MEK/PI3K inhibitors
23	Reaction 30 ACT [5635.852]	Reaction 18 ACT [1.117]	8.232	STEROID-PL	Akt/MEK inhibitors
24	Reaction 19 IN [1.580]	Reaction 24^b ACT [5635.829]	12.364	STEROID-PL	MEK/PI3K inhibitors
25	Reaction 26 ACT [6693.704]	Reaction 17 IN [1.804]	10.707	STEROID-PL	MEK/PI3K inhibitors
26	Reaction 26 ACT [8732.574]	Reaction 17 IN [1.901]	10.788	STEROID-PL	MEK/PI3K inhibitors
27	Reaction 29^f IN [9843.501]	Reaction 17 IN [1.728]	6.064	STEROID-PL	Akt/MEK inhibitors
28	Reaction 29^f IN [8620.299]	Reaction 17 IN [1.711]	6.049	STEROID-PL	Akt/MEK inhibitors
29	Reaction 18 ACT [1.298]	Reaction 29^b ACT [3725.04]	6.110	STEROID-PL	Akt/MEK inhibitors
30	Reaction 20 ACT [9.343]	Reaction 30 ACT [3185.532]	8.667	STEROID-PL	Akt/MEK inhibitors
31	Reaction 26 ACT [9734.026]	Reaction 16 ACT [156.332]	18.285	STEROID-PL	MEK/PI3K inhibitors

Table 5: Target combinations in STEROID-L and STEROID-PL corresponding to combinations curated from PubMed.

fects involving target activity of less than 100 (*Min10*, Table 7). Note that large target activity (Definition 2) implies either high drug concentration or very small dissociation constant, both of which are likely to cause side effects, especially if the treatment regime requires repeated drug dosing [10]. STEROID-PL *identified a majority of the top-10 solutions (90%)*. Most of the targets identified are located downstream of the MAPK-PI3K network. Since none of our solutions corresponded to the curated combinations in PubMed (Table 3), we performed a targeted literature search for each predicted combination (Table 7) to see if it had ever been performed. We found that *60% of the solutions in Min10, all found using* STEROID-PL, *have high biological relevance as potential target combinations.* For instance, from experiments, profound growth inhibition and apoptosis were observed in CI-1040 (MEK1/2 inhibitor) treated ovarian cancer cells with mutations in KRAS or BRAF [37] and these tumors typically overexpress DUSP4 (ERK

No.	Target 1 [Activity]	Target 2 [Activity]	ρ	Approach	PubMed Combination
32	Reaction 16 ACT [173.567]	Reaction 24^b ACT [8241.738]	21.810	STEROID-SL	MEK/PI3K inhibitors
33	Reaction 30 IN [2404.32]	Reaction 17 IN [1.799]	5.505	STEROID-SL	Akt/MEK inhibitors
34	Reaction 16 ACT [183.383]	Reaction 32 IN [7604.421]	11.855	STEROID-SL	Akt/MEK inhibitors
35	Reaction 16 ACT [187.154]	Reaction 33 ACT [3016.774]	11.905	STEROID-SL	Akt/MEK inhibitors
36	Reaction 33 ACT [9475.118]	Reaction 16 ACT [176.403]	11.760	STEROID-SL	Akt/MEK inhibitors
37	Reaction 30 ACT [906.431]	Reaction 19 IN [1.452]	7.790	STEROID-SL	Akt/MEK inhibitors
38	Reaction 33 ACT [1760.639]	Reaction 17 IN [1.784]	4.192	STEROID-SL	Akt/MEK inhibitors
39	Reaction 24^b ACT [5924.411]	Reaction 16 ACT [181.304]	20.867	STEROID-ML	MEK/PI3K inhibitors
40	Reaction 24^b ACT [8227.259]	Reaction 16 ACT [159.127]	21.578	STEROID-ML	MEK/PI3K inhibitors
41	Reaction 18 ACT [1.306]	Reaction 29^f IN [5263.823]	6.119	STEROID-ML	Akt/MEK inhibitors
42	Reaction 16 ACT [146.238]	Reaction 29^b ACT [6797.494]	13.535	STEROID-ML	Akt/MEK inhibitors
43	Reaction 22 ACT [2.148]	Reaction 24^b ACT [6121.44]	12.102	STEROID-ML	MEK/PI3K inhibitors
44	Reaction 29^b ACT [5459.28]	Reaction 16 ACT [148.572]	13.575	STEROID-ML	Akt/MEK inhibitors
45	Reaction 20 ACT [13.466]	Reaction 26 ACT [9016.078]	11.550	STEROID-P	MEK/PI3K inhibitors
46	Reaction 33 ACT [7787.74]	Reaction 16 ACT [175.569]	11.748	STEROID-P	Akt/MEK inhibitors
47	Reaction 16 ACT [176.688]	Reaction 32 IN [163.357]	11.747	STEROID-P	Akt/MEK inhibitors
48	Reaction 26 ACT [970.527]	Reaction 16 ACT [157.165]	18.277	STEROID-P	MEK/PI3K inhibitors
49	Reaction 30 IN [5125.455]	Reaction 20 ACT [12.882]	6.267	STEROID-P	Akt/MEK inhibitors
50	Reaction 16 ACT [245.303]	Reaction 30 ACT [4572.657]	16.086	STEROID-P	Akt/MEK inhibitors
51	Reaction 24^b ACT [6828.242]	Reaction 16 ACT [178.781]	21.305	STEROID-P	MEK/PI3K inhibitors
52	Reaction 32 IN [5453.004]	Reaction 19 IN [1.796]	4.158	STEROID-P	Akt/MEK inhibitors
53	Reaction 16 ACT [178.010]	Reaction 24^b ACT [3238.351]	18.544	STEROID-S	MEK/PI3K inhibitors
54	Reaction 26 ACT [4330.076]	Reaction 16 ACT [150.770]	18.190	STEROID-M	MEK/PI3K inhibitors
55	Reaction 16 ACT [323.857]	Reaction 31 IN [1640.247]	15.534	MCSA	Akt/MEK inhibitors

Table 6: Target combinations in STEROID-SL, STEROID-ML, STEROID-P, STEROID-S, STEROID-M and MCSA corresponding to combinations curated from PubMed. Numbering continues from Table 5

phosphatase) [41]. This correlates with our computational prediction of combinations 1, 6 and 9 involving ERK phosphatase activator and ERK (or MEK) kinase inhibitor. Note that there is currently no known inhibitor that acts directly on ERK and ERK inhibition is typically achieved through MEK inhibitors. Furthermore, the predicted combination 4 (ERK kinase inhibitor and tyrosine kinase inhibitor) correlates with the fact that the combination of cetuximab (tyrosine kinase inhibitor [24]) and AZD6244 (MEK inhibitor) is currently undergoing a phase 1 clinical trial for solid tumors (NCT01217450). For predicted combinations 7 and 10, we did not find any supporting evidence that they have been performed, successfully or otherwise, in experiments. However, individual components of these combinations have shown efficacy in ovarian cancer [22, 36], and warrant further investigation as potential target combinations. In the remaining predicted combinations, two (combinations 2 and 3) are more akin to monotherapies than combination therapies as they involve the same type of drug (ERK or MEK kinase inhibitor) and we did not find any supporting evidence for the other two (combinations 5 and 8). *We*

No.	Target 1 [Activity]	Target 2 [Activity]	ρ	Method
1	Reaction 21 IN [1.858]	Reaction 22 ACT [0.173]	1.010	STEROID-PL
2	Reaction 21 IN [2.063]	Reaction 19 IN [0.055]	1.010	STEROID-PL
3	Reaction 21 IN [2.248]	Reaction 17 IN [$2.5{\times}10^-4$]	1.017	STEROID-PL
4	Reaction 21 IN [2.248]	Reaction 1^f IN [$3.92{\times}10^-3$]	1.022	STEROID-PL
5	Reaction 22 ACT [2.227]	Reaction 8^f IN [$1.2{\times}10^-4$]	1.023	STEROID-PL
6	Reaction 22 ACT [2.086]	Reaction 21 IN [0.064]	1.026	STEROID-PL
7	Reaction 22 ACT [2.143]	Reaction 13 IN [0.065]	1.028	STEROID-PL
8	Reaction 22 ACT [2.222]	Reaction 5^b ACT [0.011]	1.031	STEROID-L
9	Reaction 22 ACT [2.216]	Reaction 17 IN [0.009]	1.032	STEROID-PL
10	Reaction 21 IN [2.163]	Reaction 14 ACT [0.010]	1.036	STEROID-PL

Target in Table 7	Count	Reaction in [18]	Description
Reaction 21 IN	6	ERKP→ERKPP	ERK kinase inhibitor
Reaction 22 ACT	6	ERKPP→ERKP	ERK phosphatase activator
Reaction 17 IN	2	MEKP→MEKPP	MEK kinase inhibitor
Reaction 1^f IN	1	R+HRG→RHRG	tyrosine kinase inhibitor
Reaction 5^b ACT	1	RShc→RP+Shc	activator of Shc dissociation from RP
Reaction 8^f IN	1	RShGS→ShGS+RP	inhibitor of ShGS dissociation from RP
Reaction 13 IN	1	Raf→Raf*	Raf inhibitor
Reaction 14 ACT	1	Raf* →Raf	Raf inhibitor
Reaction 19 IN	1	ERK→ERKP	ERK kinase inhibitor

Table 7: Top: Target combinations in *Min-10*. Bottom: Details of targets. *Count* represents the number of occurrences of the target in the top table.

conclude that STEROID *can identify biologically relevant combinations with low off-target effects, suggesting that heuristics are useful in improving the solution quality.*

Then, we proceeded to examine target combinations with large off-target effects to compare how they differ from those with small off-target effects. We obtained 10 solutions with maximum off-target effects from the pooled solution set (*Max10*, Table 8) and performed a targeted literature search for each predicted combination. In contrast to the targets in *Min10*, targets in *Max10* are mostly located upstream of the MAPK-PI3K network. All 10 combinations contain tyrosine kinase inhibitors (TKI). Clinical studies have found that although treatments with tyrosine kinase inhibitors (TKI) sometimes produce promising results, most treatment lose their effectiveness soon due to resistance often caused by activating mutations in downstream effectors of the tyrosine kinases [42]. Apart from drug resistance, another potential issue of TKI is toxicity (*e.g.*, nephrotic syndrome) due to the disruption of multiple downstream signaling pathways of the tyrosine kinases which are involved in normal organ functioning [17]. Hence, designing combinations involving TKI requires understanding the idiosyncrasy of a patient's genome in order to select other suitable targets for the combinations which can minimize TKI-resistance. The activity of TKI in the combinations should also

No.	Target 1 [Activity]	Target 2 [Activity]	ρ	Method
1	Reaction 1^b ACT [3556.86]	Reaction 28 IN [6870.24]	3998.34	MCSA
2	Reaction 28 IN [5050.14]	Reaction 1^b ACT [3601.55]	3000.73	MCSA
3	Reaction 1^b ACT [3281.75]	Reaction 28 IN [4396.67]	2775.37	STEROID-M
4	Reaction 32 ACT [4297.78]	Reaction 1^b ACT [9192.45]	361.27	STEROID-S
5	Reaction 27 ACT [993.14]	Reaction 3^b IN [4716.54]	361.22	MCSA
6	Reaction 1^b ACT [9136.46]	Reaction 31 IN [9945.12]	359.95	MCSA
7	Reaction 1^b ACT [8992.41]	Reaction 32 ACT [2464.55]	359.59	STEROID-S
8	Reaction 31 IN [1215.39]	Reaction 1^b ACT [8406.04]	353.37	MCSA
9	Reaction 32 ACT [9962.99]	Reaction 1^b ACT [8085.70]	351.05	STEROID-S
10	Reaction 15 ACT [3809.31]	Reaction 1^f IN [1415.88]	349.73	MCSA

Target	Count	Reaction in [18]	Description
Reaction 1^b ACT	8	RHRG→R+HRG	tyrosine kinase inhibitor
Reaction 28 IN	3	PIP3→PI	PIP3 phosphatase inhibitor
Reaction 32 ACT	3	AktPIP→AktPIPP	PI3 kinase activator
Reaction 31 IN	2	AktPIP→AktPIP3	PI3 kinase inhibitor
Reaction 1^f IN	1	R+HRG→RHRG	tyrosine kinase inhibitor
Reaction 3^b ACT	1	RP→RHRG2	tyrosine kinase inhibitor
Reaction 15 ACT	1	MEK→MEKP	MEK kinase activator
Reaction 27 ACT	1	PI→PIP3	PI3 kinase activator

Table 8: Top: Target combinations in *Max-10*. Bottom: Details of targets. *Count* represents the number of occurrences of the target in the top table.

be kept low to reduce potential toxicity. However, the predicted combinations involve TKI at high activity level, making them less than ideal. In addition, we noted that several of the targets identified in the predicted combinations (50%) involve promoters of known mediators of cancer (*e.g.*, MEK kinase [39]). Such target combinations, though counter-intuitive, can still achieve the therapeutic goal if the effect of other targets can offset the pro-cancer signals, resulting in an overall anti-cancer signal. However, extra caution should be exercised for these combinations since improper management of the balance between the pro- and anti-cancer signal can easily aggravate the cancer. Hence, *target combinations with predicted large off-target effects may be indicative of less effective and/or more toxic combinations in the real world and correlate well with the notion that large off-target effects suggest indication of toxicity. These set of experiments give us a successful "negative control" which serve further to validate the feasibility of using the literature search method for evaluating the target combinations.*

5.2 Runtime Performance

In this set of experiments, we analyzed the execution time needed to complete analysis for different approaches. Fig. 3 plots the results. We can make two key observations. First, our proposed heuristic approach is *an order of magnitude faster than state-of-the-art techniques* (MCSA). Second, approaches incorporat-
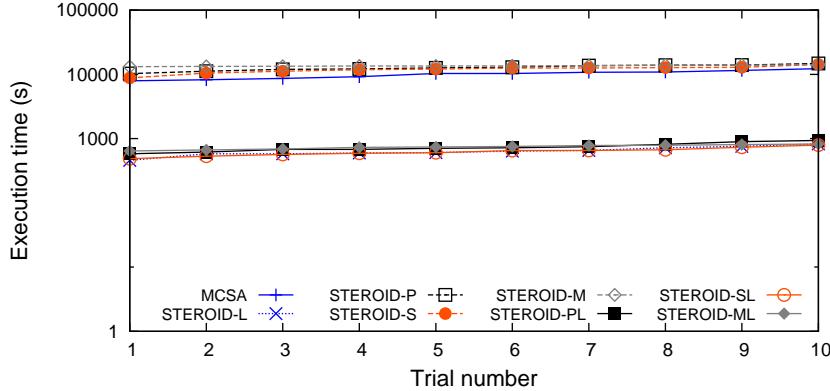
Figure 3: Runtime performance.

ing LOEWE are approximately an order of magnitude faster than other approaches, likely because LOEWE-heuristics avoid doing full evaluation on non-synergistic combinations which are excluded during candidate generation.

5.3 Off-Target Effects

Next, we studied the effects of using different approaches on the off-target effects (Section 3.2) by comparing various descriptive statistics (Fig. 4). We also pooled together the solutions obtained from the 10 trials (pooled trial) and compared the cumulative distribution functions (CDF) using t-test and Kolmogorov-Smirnov (KS) test for further analysis. Observe that STEROID-PL achieved the lowest minimum, maximum, average and median off-target effects. For the pooled trials, using one-tailed t-test, STEROID-PL produced solutions with lower off-target effects compared with other approaches ($p<0.05$); for the CDF, using one-tailed KS-test, STEROID-PL produced CDF which lie further to the left compared with other approaches ($p<0.005$). This implies that STEROID-PL *produces solutions with significantly lower off-target effects when compared to other approaches.* This could be due to PANI and LOEWE seeking solutions with reduced off-target effects in a complementary manner: LOEWE selects for lower target activity by enforcing target synergy while PANI selects for targets with lower off-target effects at higher probability.

5.4 Combination Characterization

In this set of experiments, we characterized the solutions based on the target interaction (synergistic, additive or antagonistic) and the combination type (activators, inhibitors, or mixed activator and inhibitor). *Approaches incorporating* LOEWE *produced synergistic combinations by default while other approaches produce mainly antagonistic combinations* (Fig. 5a). In terms of the combination type (Fig. 5b), we observed that *the bulk of the combinations found are mixed activator and inhibitor.* Development of inhibitors of protein-protein interaction are per-
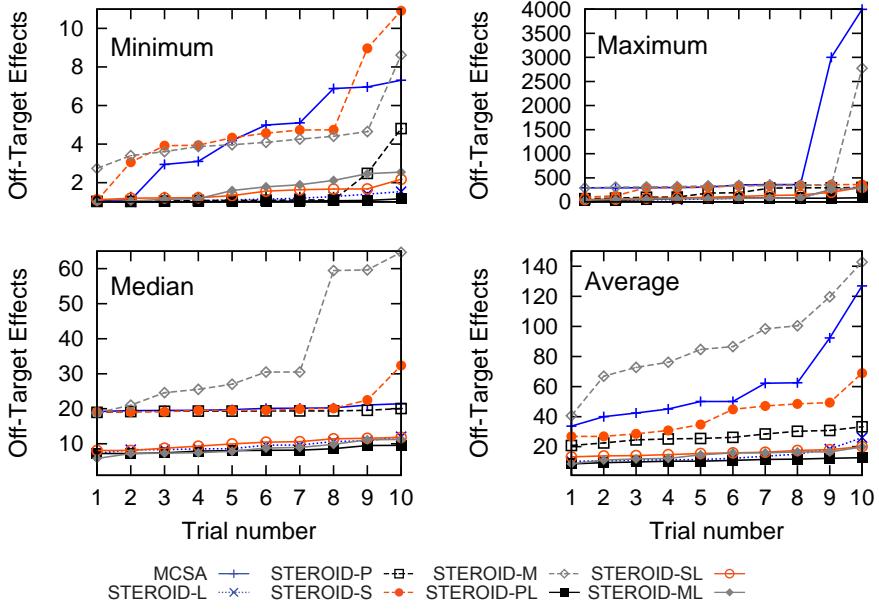
Figure 4: Off-target effects.
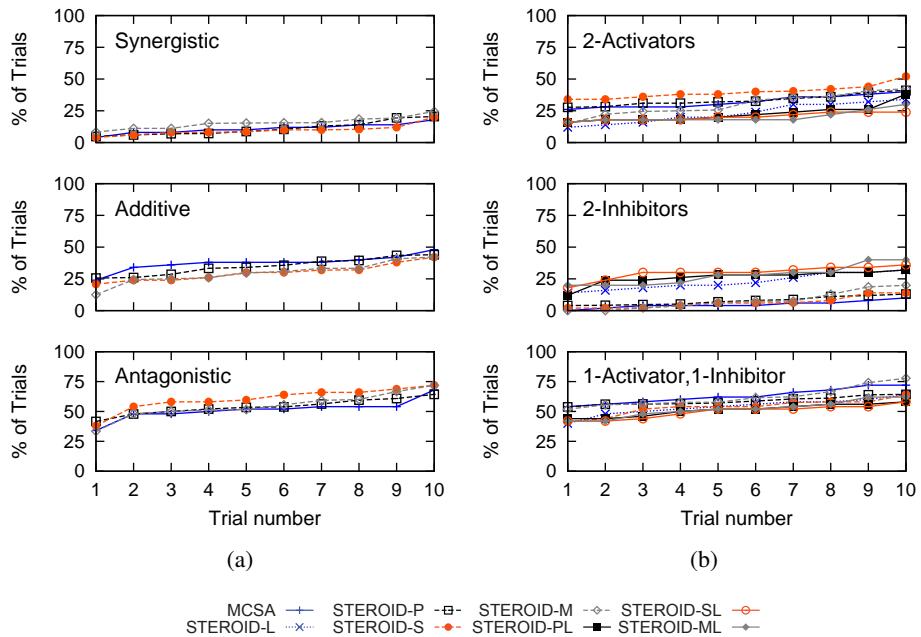
Figure 5: (a) Target interaction. (b) Combination types.
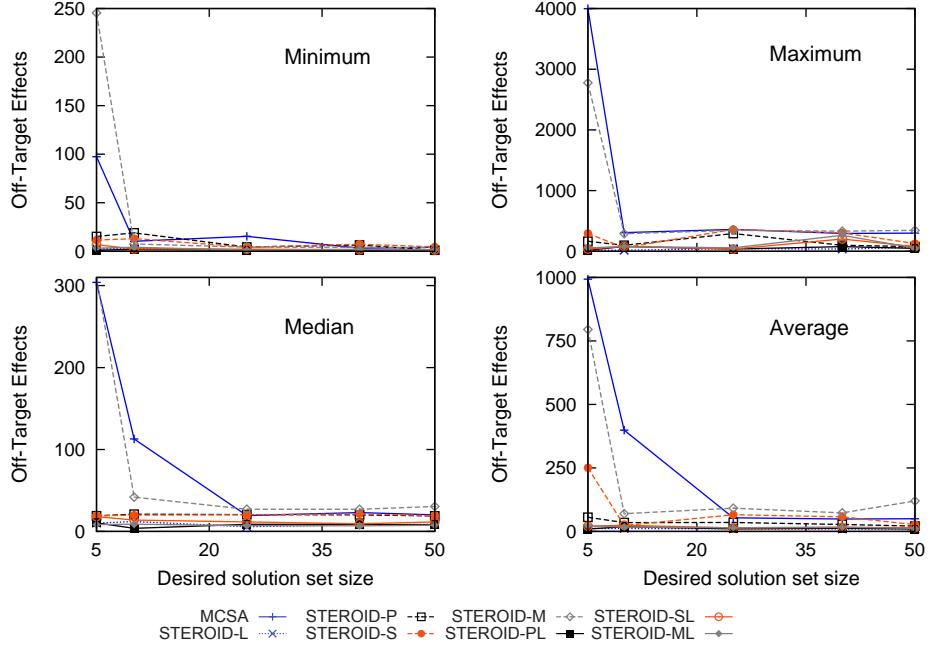
Figure 6: Effect of $\mathcal{N}$ on off-target effects.

ceived to be easier than activators which have to achieve binding and be a good replication of the protein interaction to stimulate increase in activity [16]. *Incorporating* LOEWE-*based heuristics improved the fraction of trials with 2-inhibitors by about 1 to 3 folds*.

5.5 Effect of Parameters

Here, we studied the effects of using different parameter values on the results. These parameters are namely, the solution set size $\mathcal{N}$; initial temperature τ_0; limits on iterations per cycle i_{max}; adjustment factor for therapeutic effect θ_t; adjustment factor for target interaction θ_a; and selective pressure λ^+. The parameters were varied one at a time. Unless otherwise stated, the parameters were set to default values.

Effect of $\mathcal{N}$. First, we investigated the effect of $\mathcal{N}$ by varying it in the range of $\{5,10,25,40,50\}$. We observed from Fig. 6 that the off-target effects generally decreased as $\mathcal{N}$ increased, and converged at $\mathcal{N} > 30$. The increase in $\mathcal{N}$ demanded longer execution time to complete analysis as shown in Fig. 7. We set $\mathcal{N} = 50$ for subsequent experiments since all the methods were able to identify ~ 50 solutions (Fig. 7).

Effect of τ_0. Second, we examined the effect of varying τ_0 in the range of $\{5,25,50,75,100\}$. We observed from Fig. 8 that the off-target effects generally decreased as τ_0 increased. In particular, the minimum off-target effects converged at

25
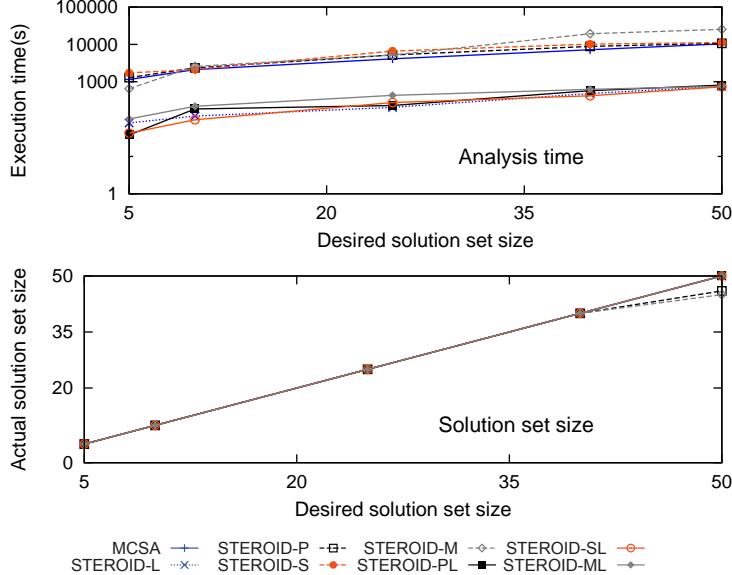
Figure 7: Effect of $\mathcal{N}$ on execution time and actual solution size.

$\tau_0 \geq 75$. From Fig. 9, we noted that variation in τ_0 had little effect on the execution time needed to complete analysis for STEROID-L, STEROID-PL, STEROID-SL and STEROID-ML. For other approaches, execution time increased as τ_0 is increased. In terms of the actual solution set size found, increasing τ_0 led to an increase in the number of solutions found, especially for approaches that did not implement LOEWE-based heuristics. These observations imply that the LOEWE-based heuristics enriches the search space significantly, allowing the required number of solutions to be found in a much shorter time. For subsequent experiments, we set $\tau_0 = 100$ since all the methods were able to identify ~ 50 solutions (Fig. 9).

Effect of i_{max}. Third, we investigated the effect of varying i_{max} in the range of $\{5, 100, 250, 350, 500\}$. Similar to the results mentioned above, we observed from Fig. 10 that the minimum off-target effects decreased as i_{max} increased. Convergence is achieved at $i_{max} \geq 350$. No visible trends were observed for the maximum, median and average off-target effects. From Fig. 11, we noted that variation in i_{max} had little effect on the execution time needed to complete analysis when $i_{max} \geq 100$ for STEROID-L, STEROID-PL, STEROID-SL and STEROID-ML. For other approaches, execution time increased as i_{max} is increased. In terms of the actual solution set size found, increasing i_{max} generally led to an increase in the number of solutions found, especially for approaches that did not implement LOEWE-based heuristics. These observations imply that the LOEWE-based heuristics enriches the search space significantly, allowing the required number of solutions to be found in a much shorter time. For subsequent experiments, we set $i_{max} = 500$ since all the methods were able to identify ~ 50 solutions (Fig. 9).

Effect of θ_t. Fourth, we examined the effect of varying θ_t in the range of

26
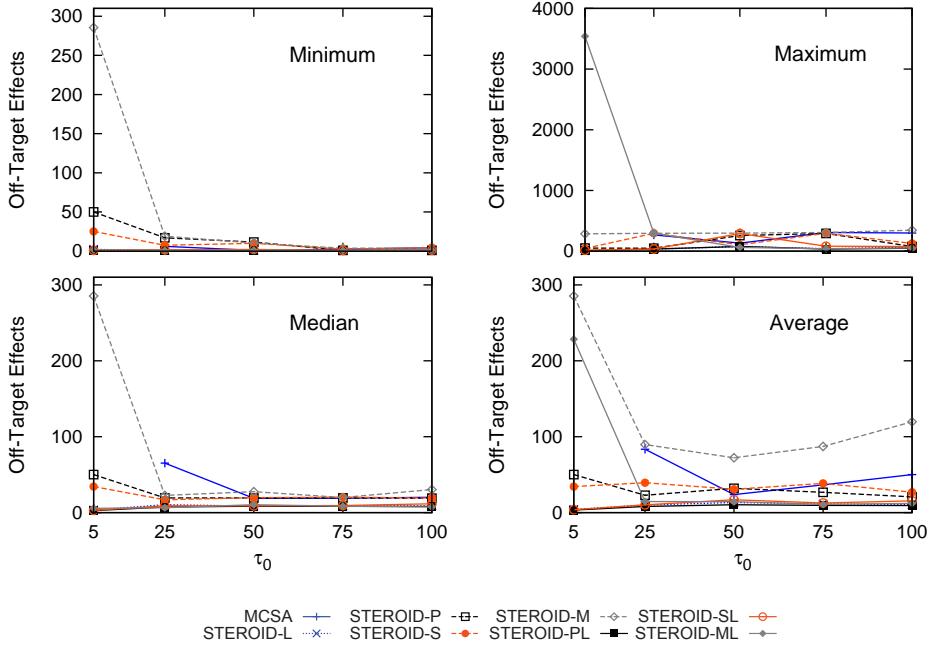
Figure 8: Effect of τ_0 on off-target effects.
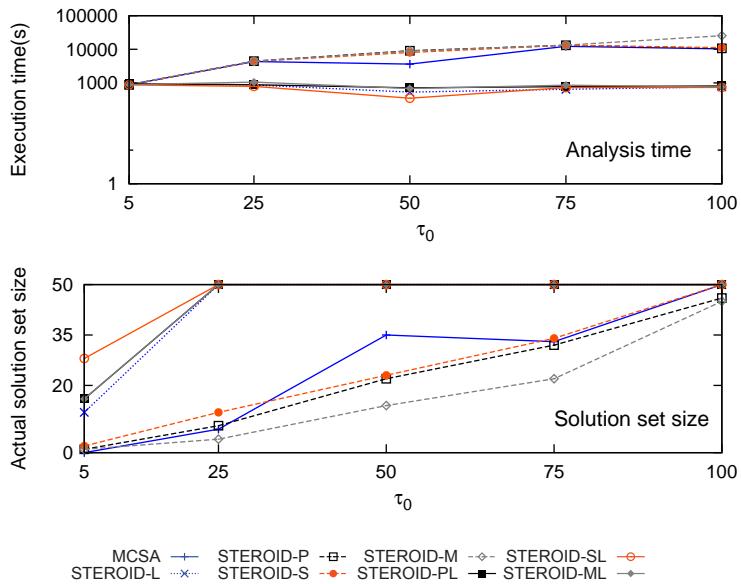
Figure 9: Effect of τ_0 on execution time and actual solution size.
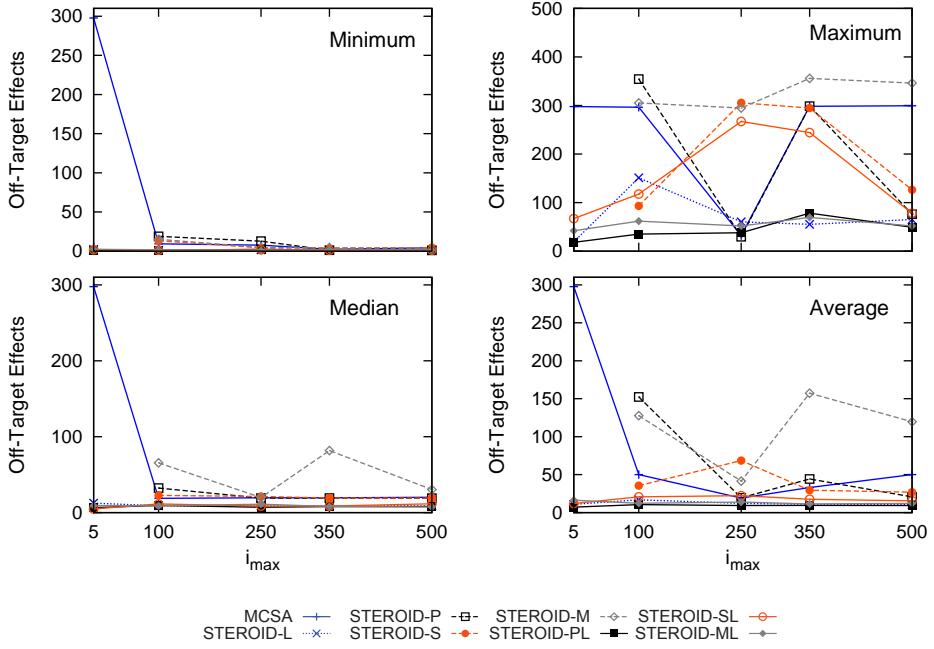
Figure 10: Effect of i_{max} on off-target effects.
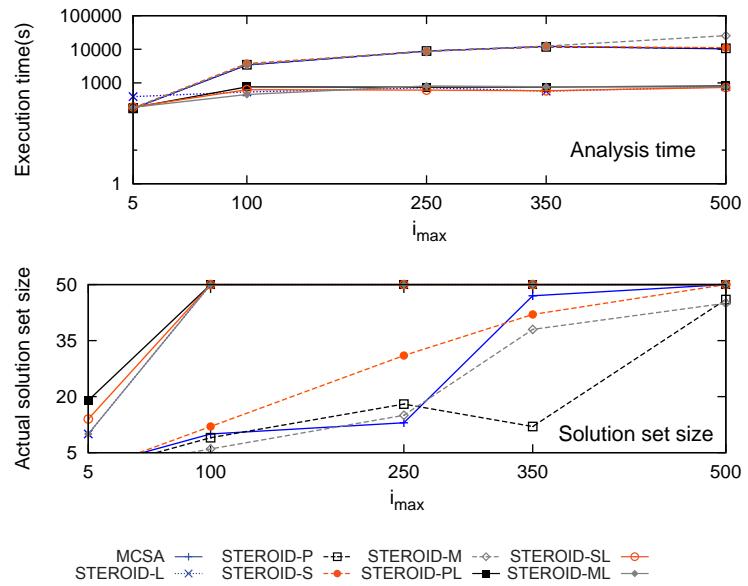
Figure 11: Effect of i_{max} on execution time and actual solution size.
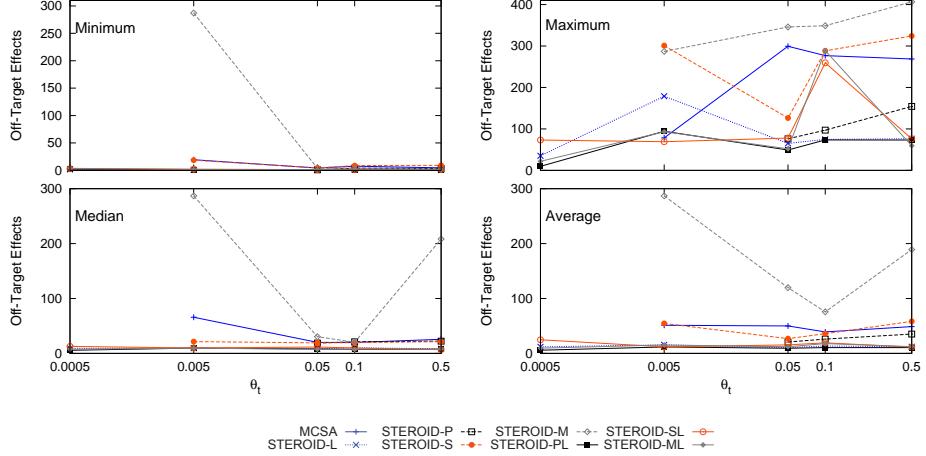
Figure 12: Effect of θ_t on off-target effects.

{0.0005,0.005, 0.05,0.1,0.5}. The more prominent trend in Fig. 12 is a decrease in the minimum off-target effects as θ_t increases. From Fig. 13, we also observed that as θ_t increased, the actual solution set size increased while the execution time decreased. This is due to further relaxation of the condition for therapeutic effect which allowed more candidates to be accepted, which also meant that there is greater likelihood to find a smaller minimum off-target effect from amongst the candidates. However, this is achieved at the expense of solutions moving further away from the desired therapeutic effect (Fig. 14 and Fig. 15). We set $\theta_t = 0.05$ for subsequent experiments since solutions were within a close range to the desired therapeutic effect and a reasonable number of solutions could be found for all the approaches in the analysis. Interestingly, we noted that compared to other approaches, the minimum, maximum, median and average off-target effects of STEROID-PL were robust to changes in θ_t.

Effect of θ_a. Fifth, we examined the effect of varying θ_a in the range of {0.0005,0.005, 0.05,0.1,0.5}. The off-target effects had no visible trends when θ_a was varied (Fig. 16). There was generally a reduction in actual solution set size when θ_a was increased, probably due to more stringent condition for target synergism resulting in fewer candidates satisfying the target synergism condition (Section 4.1) in LOEWE-based heuristic. This resulted in a general increase in execution time as reflected in Fig. 17 since more candidates have to be considered before suitable solutions are found. This is especially so for approaches incorporating LOEWE-based heuristics when $\theta_a > 0.05$. We set $\theta_a = 0.05$ for subsequent experiments since majority of the approaches could still find a reasonable number of solutions. Interestingly, we noted that compared to other approaches, the minimum, maximum, median and average off-target effects of STEROID-PL were robust to changes in θ_a.

**Effect of λ^+.** Finally, we investigated the effect of varying λ^+ in the range of
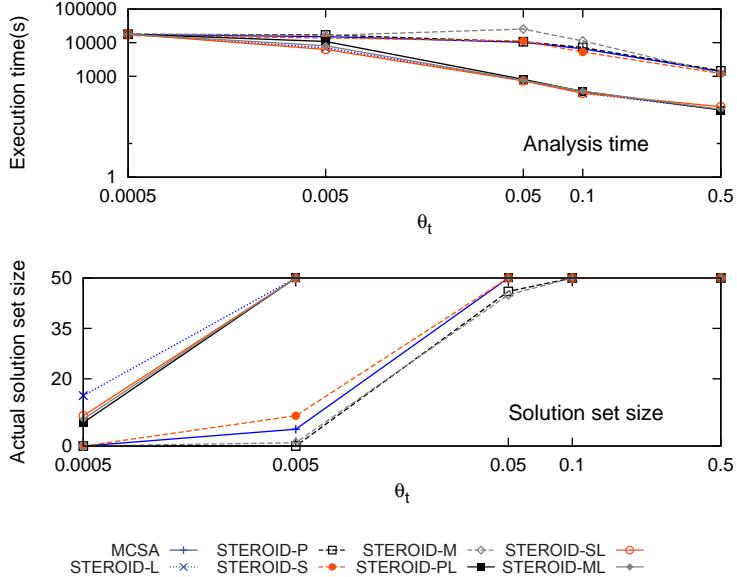
Figure 13: Effect of θ_t on execution time and actual solution size.

{1,1.3,1.5,1.8,2}. Recall from Section 4.1 that λ^+ is a parameter used to control the expected sampling rate of individual target in the target prioritization-based heuristics. Hence, varying λ^+ is likely to affect only the results of STEROID-P, STEROID-S, STEROID-M, STEROID-PL, STEROID-SL and STEROID-ML. Fig. 18 shows the effect of varying λ^+ on the off-target effects for these approaches. We observed that there was a noticeable trend of decreasing average and median off-target effects as λ^+ was increased for STEROID-P and STEROID-S. For STEROID-M, a reverse trend was observed. Approaches incorporating LOEWE-based heuristics were unaffected by λ^+ variation and were lower in terms of off-target effects in general. These observations suggest that (1) targets ranked higher by PANI and LSA tend to yield combinations with lower off-target effects since larger λ^+ implies a higher probability of selecting higher ranked individual target. The reverse is true for MPSA; (2) LOEWE-based heuristics which operates independently of λ^+ improve the robustness of STEROID. There was generally an increase in actual solution set size when λ^+ was increased for STEROID-P, STEROID-S and STEROID-M, which suggest that higher ranked individual targets were more likely to yield target combinations satisfying the user-specified therapeutic condition. We set $\lambda^+ = 1.8$ since the solution set size peaked at that value for all approaches.

Target Combinations of Different Sizes. Finally, we investigated the effects of combination size $|\mathcal{S}|$ on the result. From previous experiments, we noted that STEROID-PL produced the best results in terms of execution time and off-target effects. Hence, we shall use STEROID-PL in this set of experiments and vary $|\mathcal{S}|$ in the range of {2,3,4,5,6}. From Fig. 20, we observed that the off-target effects increased as the combination size increased. This is expected since incorporat-
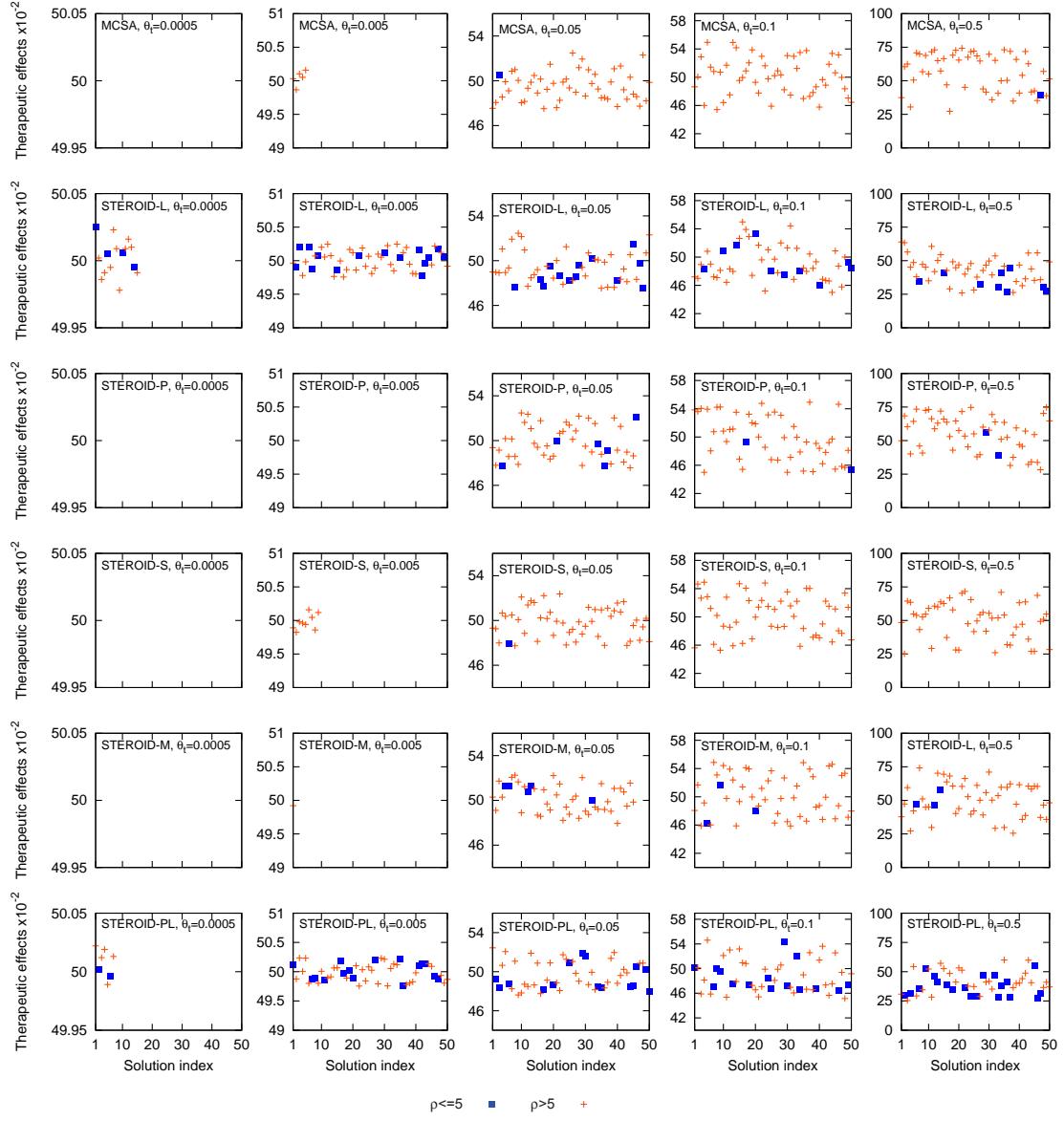
Figure 14: Solution space obtained when θ_t is varied for MCSA, STEROID-L, STEROID-P, STEROID-S, STEROID-M and STEROID-PL.
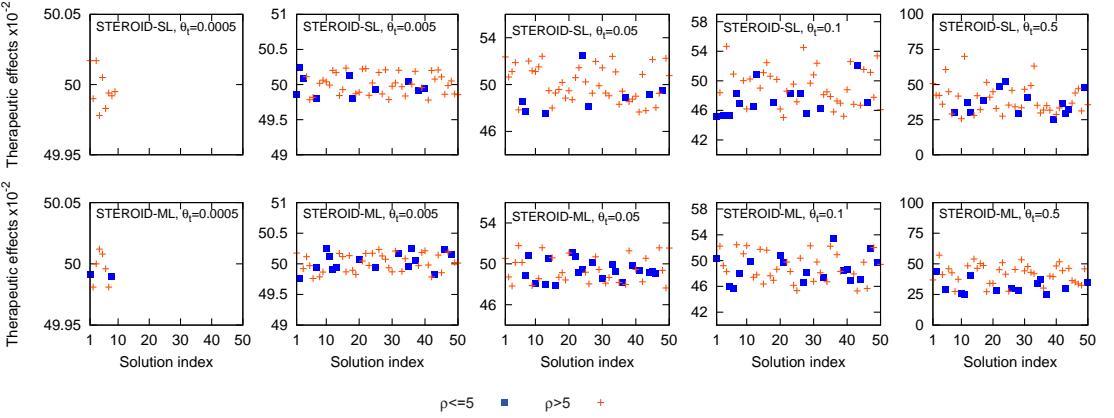
Figure 15: Solution space obtained when θ_t is varied for STEROID-SL and STEROID-ML.

ing additional targets into the combination is likely to cause perturbation of more number of downstream nodes of these targets which inevitably increases the off-target effects. The increase in combination size also resulted in a slight increase in execution time as observed in Fig. 21 and this is probably due to additional computation required for the selection of additional targets and their activities. Tables 9 to 11 list the top-5 results for the entire range of $|\mathcal{S}|$ we tested. We noted that targets closer to the output node are selected for the combinations since all the listed combinations include either a ERK (or MEK) kinase inhibitor or a ERK (or MEK) phosphatase activator. This is expected since PANI tend to prioritize these targets over those that are further from the output node [8]. We also observed that as $|\mathcal{S}|$ increased, there was a tendency to include less desirable targets (*i.e.*, promoter of pro-cancer signals) such as tyrosine kinase activator. Hence, additional rules such as exclusion of pro-cancer signal promoters should be considered to yield better target combinations when combinations of larger size are desired.

6 Conclusions & Future Work

In this work, we describe STEROID, the first heuristic approach based on LOEWE and target prioritization for modifying target combinations in a signaling network using the simulated annealing framework. Our results highlight the importance of using heuristics to improve the process of generating appropriate candidates during the search for target combinations in terms of both execution time and result quality. STEROID-PL, particularly, produces superior results with high biological relevance and significantly lower off-target results. It also discovers potential combinations (*e.g.*, ERK phosphatase activator with Raf inhibitor) for further exploration. The good performance of STEROID-PL is likely due to the selection of synergistic targets (LOEWE-based heuristic) based on structural and kinetic properties of the network (PANI-based heuristic). Furthermore, we note that LOEWE-based heuris-
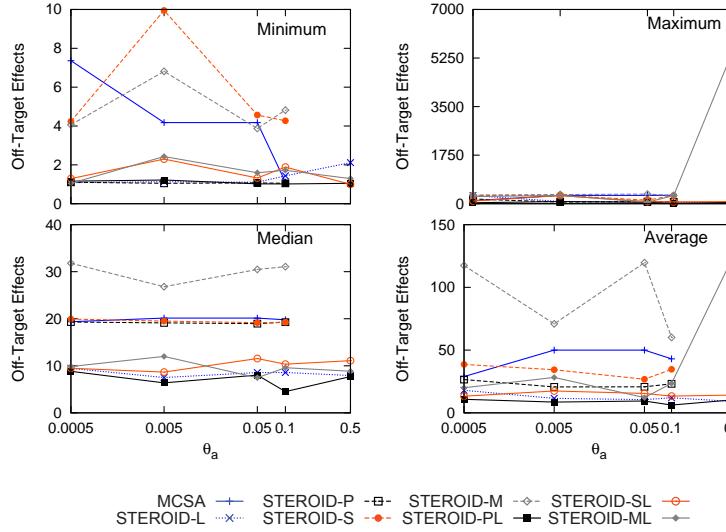
Figure 16: Effect of θ_a on off-target effects.
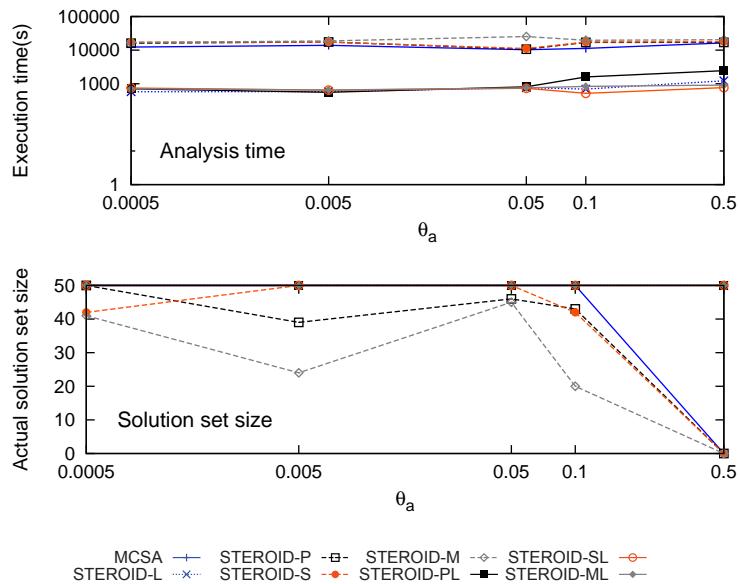
Figure 17: Effect of θ_a on execution time and actual solution size.
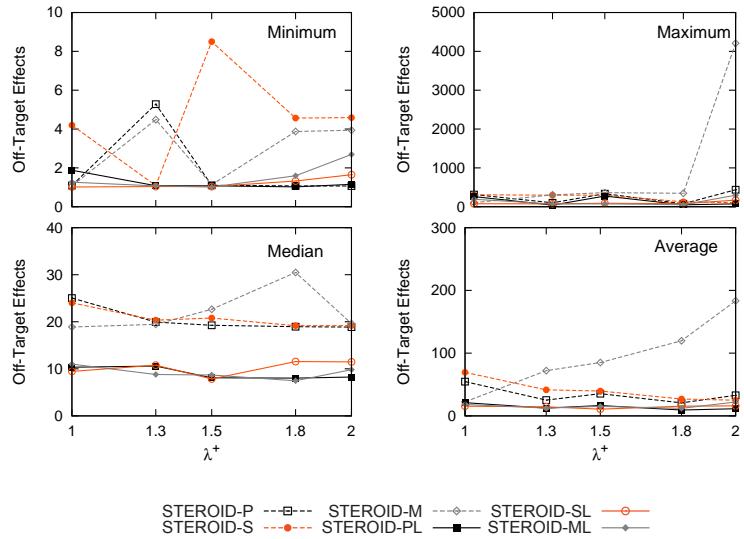
Figure 18: Effect of λ^+ on off-target effects.
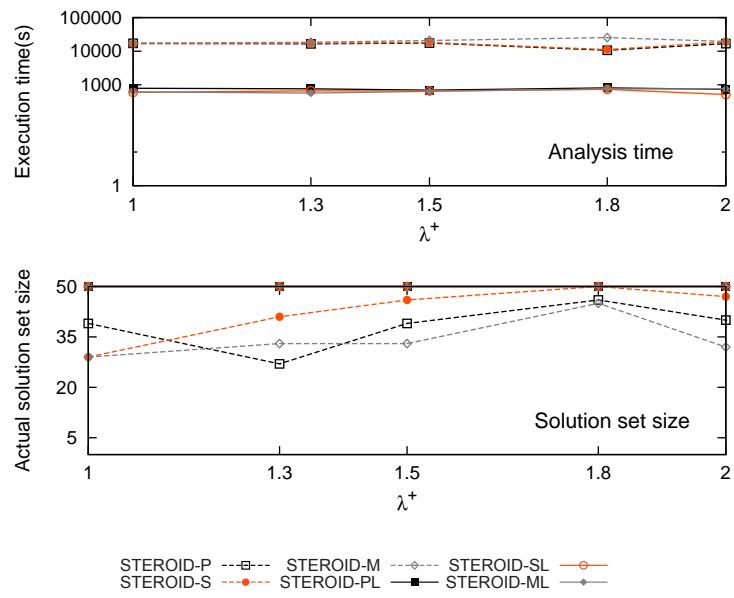
Figure 19: Effect of λ^+ on execution time and actual solution size.
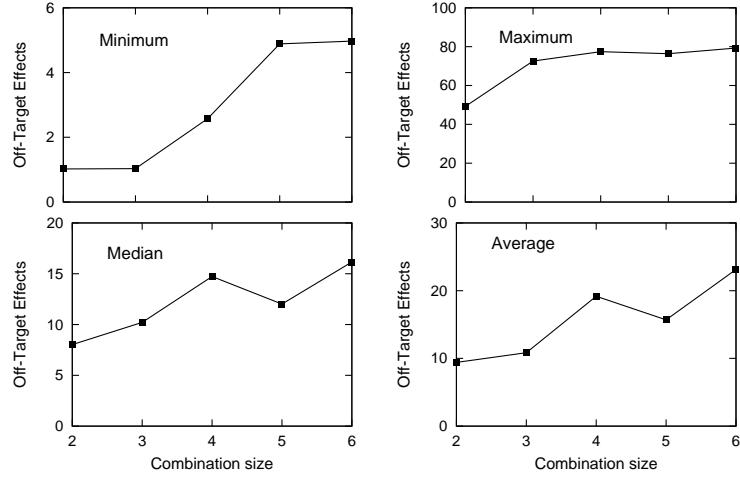
Figure 20: Effect of combination size on off-target effects for STEROID-PL.
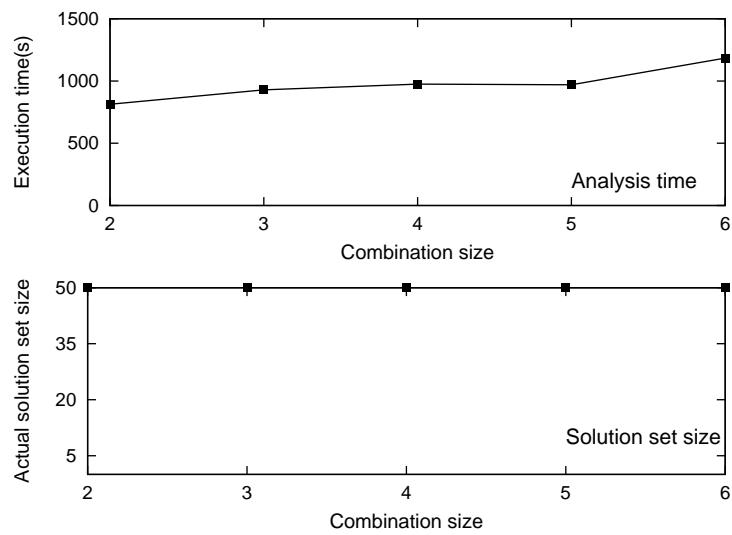
Figure 21: Effect of combination size on execution time and actual solution size for STEROID-PL.

35

|𝒮|=2

No.	Target [Activity]	Reaction in [18]	Description	ρ
1	Reaction 21 IN [2.248]	ERKP→ERKPP	ERK kinase inhibitor	1.02187
	Reaction 1^f IN [0.004]	R+HRG→RHRG	tyrosine kinase inhibitor	
2	Reaction 22 ACT [2.216]	ERKPP→ERKP	ERK phosphatase activator	1.03207
	Reaction 17 IN [0.009]	MEKP→MEKPP	MEK kinase inhibitor	
3	Reaction 22 ACT [2.121]	ERKPP→ERKP	ERK phosphatase activator	1.08848
	Reaction 14 ACT [0.023]	Raf*→Raf	Raf inhibitor	
4	Reaction 20 ACT [1.791]	ERKP→ERK	ERK phosphatase activator	1.25864
	Reaction 21 IN [1.746]	ERKP→ERKPP	ERK kinase inhibitor	
5	Reaction 19 IN [1.423]	ERK→ERKP	ERK kinase inhibitor	1.84469
	Reaction 17 IN [0.265]	MEKP→MEKPP	MEK kinase inhibitor	

|𝒮|=3

No.	Target [Activity]	Reaction in [18]	Description	ρ
1	Reaction 34^b IN [1444.745]	internalization→RP	tyrosine kinase inhibitor	1.03054
	Reaction 21 IN [0.610]	ERKP→ERKPP	ERK kinase inhibitor	
	Reaction 22 ACT [1.186]	ERKPP→ERKP	ERK phosphatase activator	
2	Reaction 21 IN [2.099]	ERKP→ERKPP	ERK kinase inhibitor	1.13798
	Reaction 19 IN [0.13054]	ERK→ERKP	ERK kinase inhibitor	
	Reaction 17 IN [0.035]	MEKP→MEKPP	MEK kinase inhibitor	
3	Reaction 21 IN [2.025]	ERKP→ERKPP	ERK kinase inhibitor	1.27025
	Reaction 20 ACT [0.820]	ERKP→ERK	ERK phosphatase activator	
	Reaction 15 IN [0.150]	MEK→MEKP	MEK kinase inhibitor	
4	Reaction 14 ACT [0.053]	Raf*→Raf	Raf inhibitor	1.79488
	Reaction 19 IN [1.551]	ERK→ERKP	ERK kinase inhibitor	
	Reaction 3^b IN [8624.076]	RP→RHRG2	tyrosine kinase activator	
5	Reaction 12 ACT [0.029]	RasGTP→RasGDP	RasGAP activator	1.90132
	Reaction 22 ACT [1.622]	ERKPP→ERKP	ERK phosphatase activator	
	Reaction 14 ACT [0.154]	Raf*→Raf	Raf inhibitor	

|𝒮|=4

No.	Target [Activity]	Reaction in [18]	Description	ρ
1	Reaction 21 IN [1.711]	ERKP→ERKPP	ERK kinase inhibitor	2.57488
	Reaction 6^f IN [0.741]	RShc→RShP	Shc kinase inhibitor	
	Reaction 8^b ACT [0.034]	ShGS+RP→RShGS	promoter of ShGS and RP binding	
	Reaction 10 IN [5927.404]	ShP→Shc	Shc phosphatase inhibitor	
2	Reaction 24^f ACT [9096.206]	RPI3K→RPI3K*	PI3K activator	4.60078
	Reaction 7^b ACT [1.140]	RShGS→RShP+GS	promoter of RShGS dissociation	
	Reaction 22 ACT [1.148]	ERKPP→ERKP	ERK phosphatase activator	
	Reaction 17 IN [0.155]	MEKP→MEKPP	MEK kinase inhibitor	
3	Reaction 31 ACT [7516.839]	AktPIP→AktPIP3	PIP kinase activator	5.85213
	Reaction 6^b ACT [1.291]	RShP→RShc	Shc phosphatase activator	
	Reaction 21 IN [1.323]	ERKP→ERKPP	ERK kinase inhibitor	
	Reaction 13 IN [0.954]	Raf→Raf*	Raf inhibitor	
4	Reaction 31 IN [2943.773]	AktPIP→AktPIP3	PIP kinase inhibitor	6.03896
	Reaction 18 ACT [1.249]	MEKPP→MEKP	MEK phosphatase activator	
	Reaction 29^f IN [3705.561]	PIP3+Akt→AktPIP3	Akt inhibitor	
	Reaction 21 IN [0.059]	ERKP→ERKPP	ERK kinase inhibitor	
5	Reaction 16 IN [3446.009]	MEKP→MEK	MEK phosphatase inhibitor	6.15524
	Reaction 14 ACT [0.671]	Raf*→Raf	Raf inhibitor	
	Reaction 5^b ACT [1.047]	RShc→RP+Shc	promoter of RShc dissociation	
	Reaction 31 ACT [2110.600]	AktPIP→AktPIP3	PIP kinase activator	

Table 9: Top-5 results of STEROID-PL for $|\mathcal{S}|=\{2,3,4\}$.

|S|=5

No.	Target [Activity]	Reaction in [18]	Description	ρ
1	Reaction 21 IN [1.299]	ERKP→ERKPP	ERK kinase inhibitor	4.88571
	Reaction 20 IN [9766.701]	ERKP→ERK	ERK phosphatase inhibitor	
	Reaction 30 IN [5523.830]	AktPIP3→AktPIP	PIP3 phosphatase inhibitor	
	Reaction 33 IN [6767.722]	AktPIPP→AktPIP	PIP3 phosphatase inhibitor	
	Reaction 18 ACT [0.520]	MEKPP→MEKP	MEK phosphatase activator	
2	Reaction 28 ACT [7302.668]	PIP3→PI	PIP3 phosphatase activator	6.39667
	Reaction 21 IN [0.705]	ERKP→ERKPP	ERK kinase inhibitor	
	Reaction 19 IN [0.964]	ERK→ERKP	ERK kinase inhibitor	
	Reaction 33 ACT [2569.067]	AktPIPP→AktPIP	PIP3 phosphatase activator	
	Reaction 3[b] IN [4039.982]	RP→RHRG2	tyrosine kinase activator	
3	Reaction 32 IN [7925.339]	AktPIP→AktPIPP	PIP kinase inhibitor	6.48116
	Reaction 12 ACT [0.040]	RasGTP→RasGDP	RasGAP activator	
	Reaction 18 ACT [0.246]	MEKPP→MEKP	MEK phosphatase activator	
	Reaction 9[b] ACT [1393.884]	GS+ShP→ShGS	promoter of GS and ShP binding	
	Reaction 8[b] ACT [0.765]	ShGS+RP→RShGS	promoter of ShGS and RP binding	
4	Reaction 12 ACT [0.237]	RasGTP→RasGDP	RasGAP activator	6.64748
	Reaction 22 ACT [0.244]	ERKPP→ERKP	ERK phosphatase activator	
	Reaction 24[b] IN [1816.184]	RPI3K*→RPI3K	PI3K activator	
	Reaction 29[f] IN [9089.066]	PIP3+Akt→AktPIP3	Akt inhibitor	
	Reaction 31 ACT [7188.479]	AktPIP→AktPIP3	PIP kinase activator	
5	Reaction 27 IN [7608.703]	PI→PIP3	PI kinase inhibitor	6.74592
	Reaction 9[b] ACT [2710.060]	GS+ShP→ShGS	promoter of GS and ShP binding	
	Reaction 10 IN [3073.074]	ShP→Shc	Shc phosphatase inhibitor	
	Reaction 13 IN [3.837]	Raf→Raf*	Raf inhibitor	
	Reaction 21 IN [0.901]	ERKP→ERKPP	ERK kinase inhibitor	

Table 10: Top-5 results of STEROID-PL for |S|=5.

tic enriches the search space, improving runtime performance by about an order of magnitude, and increasing the fraction of trials with 2-inhibitor combinations by up to 3 folds.

We note that not all target prioritization approaches improve the results and the objectives of the problem (*e.g.*, identifying therapeutic target combinations with low off-target effects) can be used to guide the selection of appropriate approaches. Further extensions of this work include incorporating various "omics" data and drug and disease information as heuristics to find target combinations that exclude combinations akin to monotherapies, and that avoid including activators of pro-disease targets as part of the combinations. In this work, we assume all nodes have equally severe off-target effects. A more flexible strategy is to implement a weighted sum off-target effects to reflect differences in severity of the off-target effects in future work. In summary, our performance comparisons demonstrate the potential value of knowledge-based heuristics for sampling and evaluating targets and STEROID provides a first step in this regard.

$|\mathcal{S}|$=6

No.	Target [Activity]	Reaction in [18]	Description	ρ
1	Reaction 25^f ACT [2019.677]	RPI3K*→RP+PI3K*	promoter of RPI3K* dissociation	4.97259
	Reaction 28 ACT [7230.289]	PIP3→PI	PIP3 phosphatase activator	
	Reaction 18 ACT [1.093]	MEKPP→MEKP	MEK phosphatase activator	
	Reaction 27 ACT [8967.207]	PI→PIP3	PI kinase activator	
	Reaction 24^b IN [4964.483]	RPI3K*→RPI3K	PI3K activator	
	Reaction 1^f IN [1.128]	R+HRG→RHRG	tyrosine kinase inhibitor	
2	Reaction 30 ACT [2380.633]	AktPIP3→AktPIP	PIP3 phosphatase activator	5.60381
	MKP3 IN [5.477]	[MKP3]$_0$	MKP3 inhibitor	
	Reaction 11 IN [13.181]	RasGDP→RasGTP	RasGEF inhibitor	
	Reaction 29^f IN [8314.225]	PIP3+Akt→AktPIP3	Akt inhibitor	
	Reaction 22 ACT [0.217]	ERKPP→ERKP	ERK phosphatase activator	
	Reaction 16 ACT [0.013]	MEKP→MEK	MEK phosphatase activator	
3	Reaction 25^b IN [2296.798]	RP+PI3K*→RPI3K*	inhibitor of RP and PI3K* binding	5.62525
	Reaction 12 ACT [0.222]	RasGTP→RasGDP	RasGAP activator	
	Reaction 16 IN [1199.750]	MEKP→MEK	MEK phosphatase inhibitor	
	Reaction 3^f IN [0.114]	RHRG2→RP	tyrosine kinase inhibitor	
	Reaction 5^f IN [0.188]	RP+Shc→RShc	inhibitor of RP and Shc binding	
	Reaction 2^b IN [2643.021]	RHRG2→2*RHRG	tyrosine kinase activator	
4	Reaction 34^b ACT [4766.160]	internalization→RP	tyrosine kinase activator	7.19757
	Reaction 18 ACT [1.268]	MEKPP→MEKP	MEK phosphatase activator	
	Reaction 33 ACT [5512.159]	AktPIPP→AktPIP	PIPP phosphatase activator	
	Reaction 11 IN [9.039]	RasGDP→RasGTP	RasGEF inhibitor	
	Reaction 27 IN [4232.607]	PI→PIP3	PI kinase inhibitor	
	Reaction 19 IN [0.011]	ERK→ERKP	ERK kinase inhibitor	
5	Reaction 27 IN [3883.773]	PI→PIP3	PI kinase inhibitor	8.52135
	Reaction 11 IN [64.156]	RasGDP→RasGTP	RasGEF inhibitor	
	Reaction 22 ACT [1.177]	ERKPP→ERKP	ERK phosphatase activator	
	Reaction 34^b ACT [7191.794]	internalization→RP	tyrosine kinase activator	
	Reaction 29^b IN [7826.257]	AktPIP3→PIP3+Akt	Akt activator	
	Reaction 8^f IN [0.152]	RShGS→ShGS+RP	inhibitor of RShGS dissociation	

Table 11: Top-5 results of STEROID-PL for $|\mathcal{S}|$=6.

References

[1] J. Avruch et al. Ras activation of the Raf kinase: Tyrosine kinase recruitment of the MAP kinase cascade. *Recent Prog Horm Res*, 56(1):127–156, 2001.

[2] R. Bast. Molecular approaches to personalizing management of ovarian cancer. *Ann. Oncol.*, 22(Suppl 8):viii5–viii15, 2011.

[3] T. Blickle et al. A comparison of selection schemes used in evolutionary algorithms. *Evol. Comput.*, 4(4):361–394, 1996.

[4] S. Brailsford et al. Constraint satisfaction problems: Algorithms and applications. *Eur. J. Oper. Res.*, 119(3):557 – 581, 1999.

[5] D. Calzolari et al. Search algorithms as a framework for the optimization of drug combinations. *PLoS Comput Biol*, 4(12):e1000249, 12 2008.

[6] W. Chiou. Critical evaluation of the potential error in pharmacokinetic studies of using the linear trapezoidal rule method for the calculation of the area under the plasma level-time curve. *J. Pharmacokinet. Pharmacodynamics*, 6:539–546, 1978.

[7] T. Chou et al. A simple generalized equation for the analysis of multiple inhibitions of michaelis-menten kinetic systems. *J. Biol. Chem.*, 252(18):6438–42, 1977.

[8] H. Chua et al. PANI:a novel algorithm for fast discovery of putative target nodes in signaling networks. www.cais.ntu.edu.sg/∼assourav/TechReports/pani-tr.pdf, 2010.

[9] H. Chua et al. PANI:a novel algorithm for fast discovery of putative target nodes in signaling networks. *ACM BCB*, 2011.

[10] R. Copeland et al. Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov*, 5(9):730–739, Sep 2006.

[11] P. Csermely et al. The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci*, 26(4):178–182, Apr 2005.

[12] B. Davies et al. AZD6244 (ARRY-142886), a potent inhibitor of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase 1/2 kinases: mechanism of action in vivo, pharmacokinetic/pharmacodynamic relationship, and potential for combination in preclinical models. *Mol. Cancer Ther.*, 6(8):2209–2219, 2007.

[13] J. Fitzgerald et al. Systems biology and combination therapy in the quest for clinical efficacy. *Nat Chem Biol*, 2(9):458–466, Sep 2006.

[14] C. Fremin et al. From basic research to clinical development of mek1/2 inhibitors for cancer therapy. *J. Hematol. Oncol.*, 3(1):8, 2010.

[15] E. Freuder et al. Systematic versus stochastic constraint satisfaction. In *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, IJCAI'95, pages 2027–2032, San Francisco, CA, USA, 1995.

[16] S. Gordo et al. Knitting and untying the protein network: modulation of protein ensembles as a therapeutic strategy. *Protein Sci*, 18(3):481–493, Mar 2009.

[17] K. Gotink et al. Anti-angiogenic tyrosine kinase inhibitors: what is their mechanism of action? *Angiogenesis*, 13(1):1–14, Mar 2010.

[18] M. Hatakeyama et al. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. *Biochem J*, 373(Pt 2):451–463, Jul 2003.

[19] D. Hu et al. Time-dependent sensitivity analysis of biological networks: Coupled MAPK and PI3K signal transduction pathways. *The Journal of Physical Chemistry A*, 110(16):5361–5370, 2006.

[20] W.-C. Hwang et al. Identification of information flow-modulating drug targets: a novel bridging paradigm for drug discovery. *Clin Pharmacol Ther*, 84(5):563–572, Nov 2008.

[21] S. Iadevaia et al. Identification of optimal drug combinations targeting cellular networks: Integrating phospho-proteomics and computational network analysis. *Cancer Res.*, 70(17):6704–6714, 2010.

[22] V. Khazak et al. Selective raf inhibition in cancer therapy. *Expert Opin Ther Targets*, 11(12):1587–1609, Dec 2007.

[23] K. Kinross et al. In vivo activity of combined pi3k/mtor and mek inhibition in a krasg12d;pten deletion mouse model of ovarian cancer. *Mol. Cancer Ther.*, 10(8):1440–1449, 2011.

[24] P. Kirkpatrick et al. Cetuximab. *Nat Rev Drug Discov*, 3:549–550, 2004.

[25] S. Klamt et al. Hypergraphs and cellular networks. *PLoS Comput Biol*, 5(5):e1000385, May 2009.

[26] Y. Koh et al. Sorafenib and mek inhibition is synergistic in medullary thyroid carcinoma in vitro. *Endocr Relat Cancer*, 19(1):29–38, Feb 2012.

[27] N. Le Novère et al. Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 34(Database issue):D689–D691, Jan 2006.

[28] F. López-Muñoz et al. Analgesic effects of combinations containing opioid drugs with either aspirin or acetaminophen in the rat. *Drug Develop Res*, 29(4):299–304, 1993.

[29] M. Los et al. *Apoptotic pathways as targets for novel therapies in cancer and other diseases*. Springer, 2005.

[30] J. Luo et al. Targeting the PI3K-Akt pathway in human cancer: Rationale and promise. *Cancer Cell*, 4(4):257 – 262, 2003.

[31] B. Markman et al. Status of pi3k inhibition and biomarker development in cancer therapeutics. *Ann Oncol*, Aug 2009.

[32] N. Metropolis et al. Equation of state calculations by fast computing machines. 21(6):1087–1092, 1953.

[33] T. Millward et al. Regulation of protein kinase cascades by protein phosphatase 2a. *Trends Biochem Sci*, 24(5):186–191, May 1999.

[34] A. Nolte et al. A note on the finite time behavior of simulated annealing. *Mathematics of Operations Research*, 25(3):pp. 476–484, 2000.

[35] T. Okuzumi et al. Inhibitor hijacking of akt activation. *Nat Chem Biol*, 5(7):484–493, Jul 2009.

[36] R. Orlowski et al. Evidence that inhibition of p44/42 mitogen-activated protein kinase signaling is a factor in proteasome inhibitor-mediated apoptosis. *J Biol Chem*, 277(31):27864–27871, 2002.

[37] G. Pohl et al. Inactivation of the mitogen-activated protein kinase pathway as a potential target-based therapy in ovarian serous tumors with kras or braf mutations. *Cancer Res*, 65(5):1994–2000, 2005.

[38] A. Reibman et al. Numerical transient analysis of markov models. *Computers & Operations Research*, 15(1):19 – 36, 1988.

[39] P. Roberts et al. Targeting the raf-mek-erk mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene*, 26(22):3291–3310, May 2007.

[40] S. Sahle et al. Simulation of biochemical networks using copasi: a complex pathway simulator. In *Proceedings of the 38th conference on Winter simulation*, WSC '06, pages 1698–1706. Winter Simulation Conference, 2006.

[41] N. Sieben et al. Differential gene expression in ovarian tumors reveals dusp 4 and serpina 5 as key regulators for benign behavior of serous borderline tumors. *J Clin Oncol*, 23(29):7257–7264, 2005.

[42] J. Sierra et al. Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy. *Mol Cancer*, 9(1):75, 2010.

[43] H. Sondermann et al. Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell*, 119(3):393–405, Oct 2004.

[44] R. Steinmetz et al. Mechanisms Regulating the Constitutive Activation of the Extracellular Signal-Regulated Kinase (ERK) Signaling Pathway in Ovarian Cancer and the Effect of Ribonucleic Acid Interference for ERK1/2 on Cancer Cell Proliferation. *Mol Endocrinol*, 18(10):2570–2582, 2004.

[45] N. van Riel. Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinform*, 7(4):364–374, 2006.

[46] D. Whitley. The genitor algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. 1989.

[47] K. Yang et al. Finding multiple target optimal intervention in disease-related molecular network. *Mol Syst Biol*, 4:228, 2008.

[48] Z. Zhang et al. Phosphorylated erk is a potential predictor of sensitivity to sorafenib when treating hepatocellular carcinoma: evidence from an in vitro study. *BMC Med*, 7:41, 2009.

[49] L. Zhao et al. Evaluation of combination chemotherapy. *Clin Cancer Res*, 10(23):7994–8004, 2004.

[50] Z. Zi et al. In silico identification of the key components and steps in IFN-gamma induced JAK-STAT signaling pathway. *FEBS Lett*, 579(5):1101–1108, Feb 2005.