Supplementary Information

Tuning Reaction Products by Constrained Optimisation

Barnaby E. Walker[a], James H. Bannock[ab], Adrian M. Nightingale[c], and John C. deMello[*ab]

[a] *Centre for Plastic Electronics, Department of Chemistry, Imperial College London, Exhibition Road, South Kensington, London SW7 2AY, UK*

[b] *Centre for Organic Electronic Materials, Department of Chemistry, NTNU, N-7491 Trondheim, Norway*

[c] *Faculty of Engineering and the Environment, University of Southampton, Southampton SO17 1BJ, UK*

[*] *Email: j.demello@imperial.ac.uk*

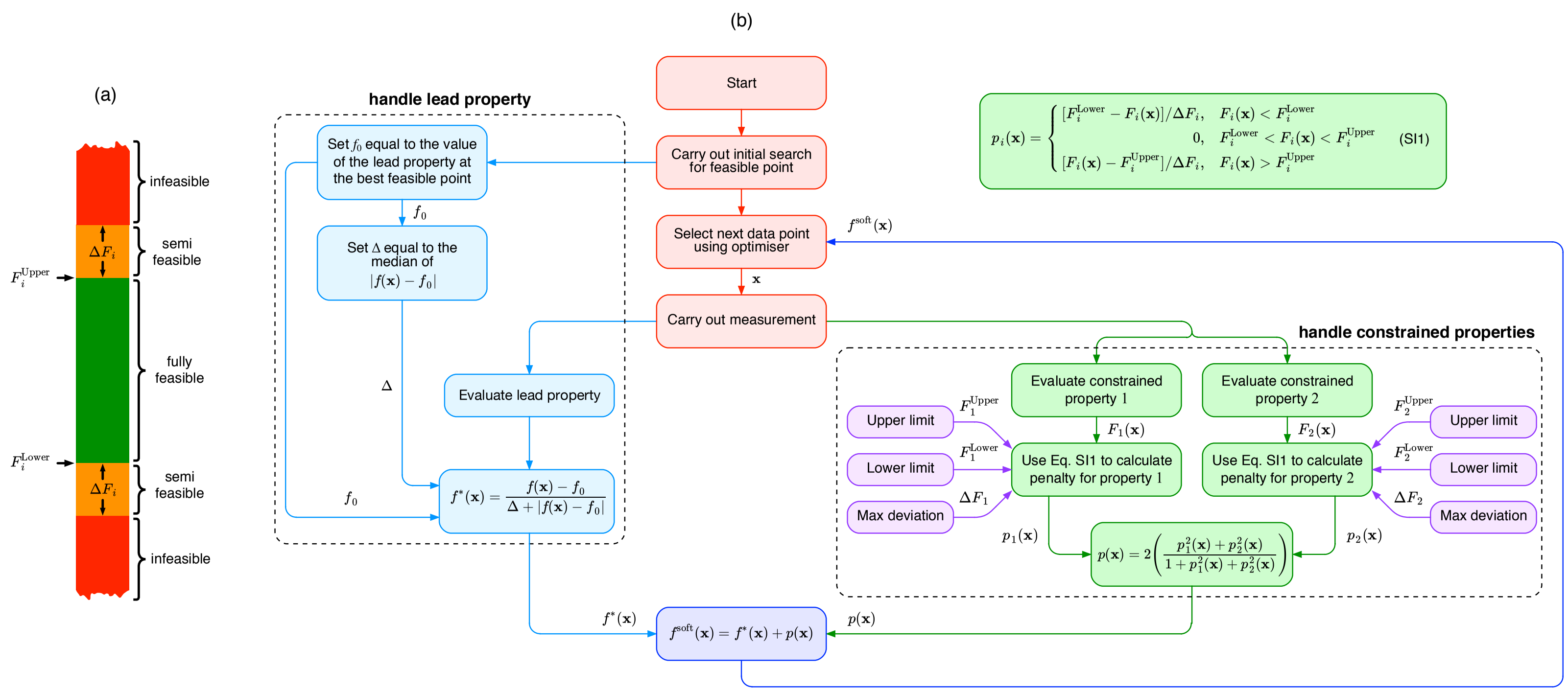

Figure S1: (a) Schematic showing the classification of property values as feasible, semi-feasible or infeasible, according to whether they lie in the preferred range $[F_i^{\mathrm{Lower}}, F_i^{\mathrm{Upper}}]$, the permitted range $[F_i^{\mathrm{Lower}} - \Delta F_i, F_i^{\mathrm{Upper}} + \Delta F_i]$, or outside the permitted range. (b) Flow chart summarising the procedure used to optimise a lead property, while at the same time enforcing two constraints. In the first phase of the optimisation, a search is carried out for a fully feasible data point that completely satisfies both constraints. Two parameters are extracted from the data acquired during this initial search: f_0, the value of the lead property at the best feasible point; and Δ, the median value of $|f(\mathbf{x}) - f_0|$ determined from all points tested in the initial search. In the second phase, a search is carried out for reaction conditions that minimise the value of the lead property, while at the same time ensuring the values of the constrained properties lie in the range $[F_i^{\mathrm{Lower}} - \Delta F_i, F_i^{\mathrm{Upper}} + \Delta F_i]$. For each tested data point x, two quantities are calculated: $f^*(\mathbf{x})$, a rescaled variant of $f(\mathbf{x})$ that lies in the range [-1,1]; and $p(\mathbf{x})$, a penalty that lies in the range [0,2]. Fully feasible data points that fully satisfy all constraints have penalty values of zero, semi-feasible data points that partially violate one or more constraints have penalty values in the range [0,1], while infeasible points that fully violate one or more constraints have penalty values in the range [1,2]. The overall merit value $f^{\mathrm{soft}}(\mathbf{x})$ is obtained by summing $f^*(\mathbf{x})$ and $p(\mathbf{x})$, which yields a value in the range [-1,3]. The constrained optimisation is carried out by searching for the location $\hat{\mathbf{x}}$ of the global minimum of $f^{\mathrm{soft}}(\mathbf{x})$.

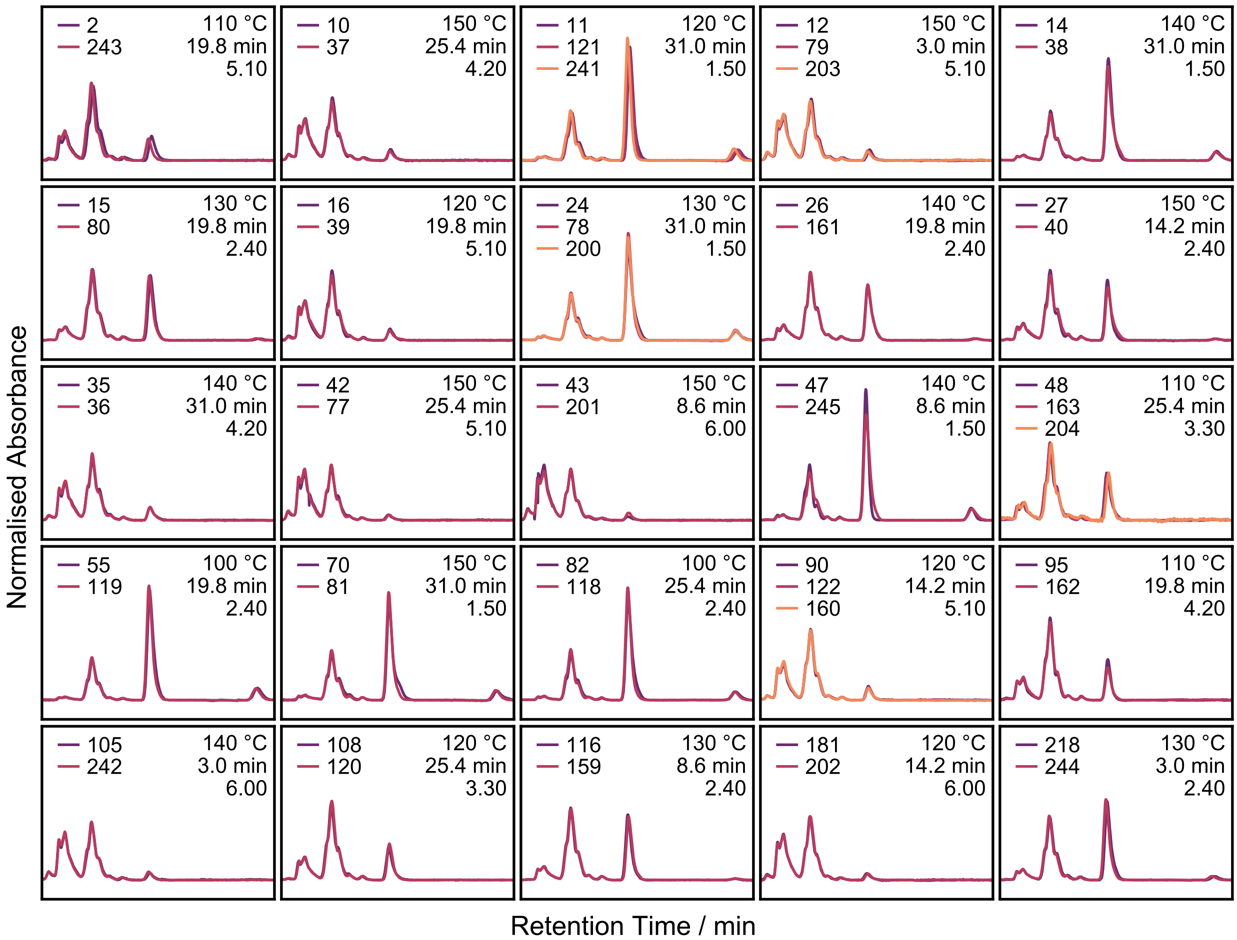

Figure S2: Graphs showing replicate chromatograms obtained during a 6 × 6 × 6 grid search of the reaction parameter space. 216 uniformly spaced grid points were tested in a randomised order, of which 191 were tested once, 20 were tested twice and 5 were tested three times. The chromatograms shown here correspond to those grid points that were tested multiple times. The numbers listed in the top left of each graph specify the measurement number of the chromatograms, while the numbers listed on the right denote the corresponding temperature, reaction time and molar ratio of sultine to C_{60}.

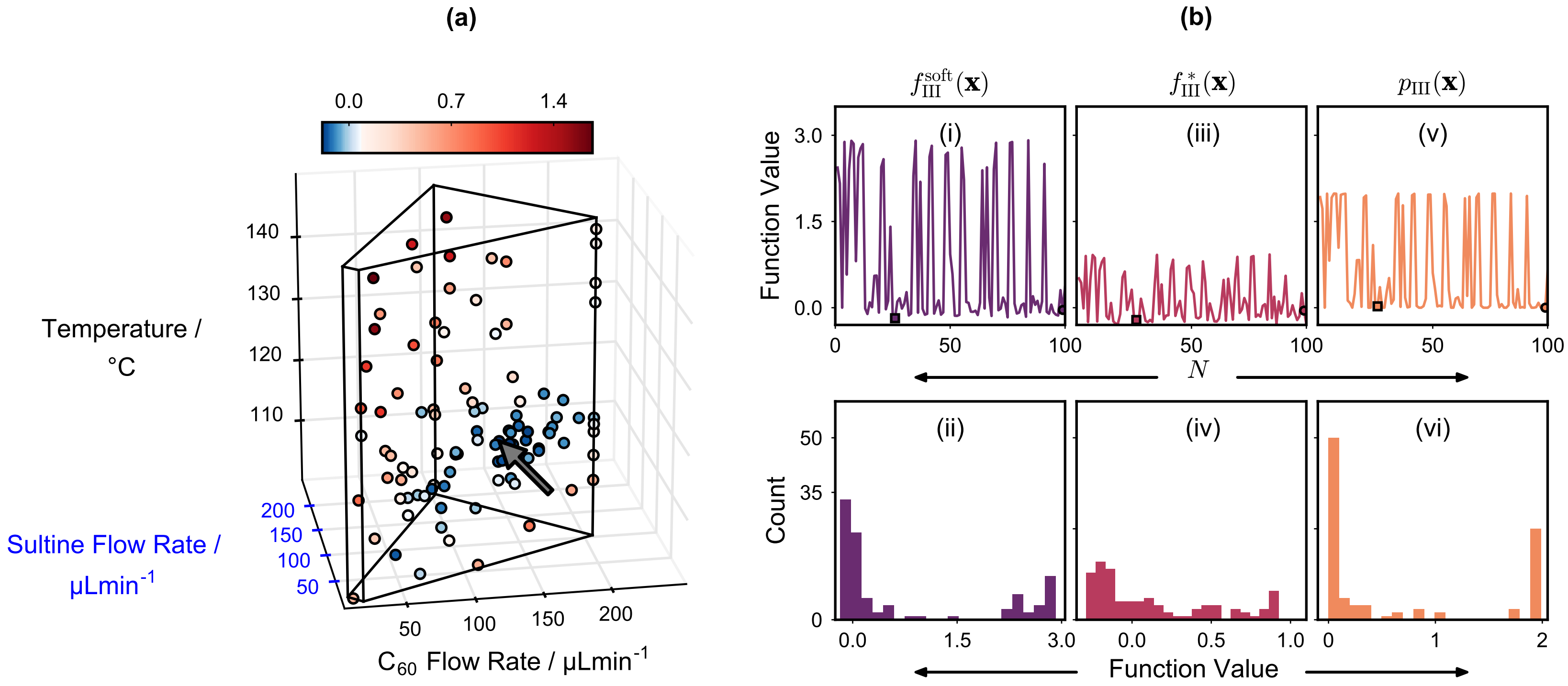

Figure S3: (a) Scatter plot for Run III, showing the influence of the sultine flow rate, the C_{60} flow rate and the temperature on the merit value $f_{\mathrm{III}}^{\mathrm{soft}}$. The location of each data point indicates the reaction conditions used, while the colour denotes the corresponding merit value. For ease of interpretation, points with merit values above the median merit value (0.087) have been coloured red, while those with merit values below the median value have been coloured blue; points with merit values close to the median value appear as white. The black cage defines the flow rate and temperature constraints. The arrows indicate the locations of the best point (i.e. the point with the lowest merit value) for the two runs. (b) Merit values for Run III expressed as a time series (i) and a histogram (ii). Also shown are the time series and histograms for $f_{\mathrm{III}}^{*}(x)$ (iii, iv) and $p_{\mathrm{III}}(x)$ (v, vi). The square and circle-shaped markers in the time-series plots correspond to the best point and the best feasible point, respectively.

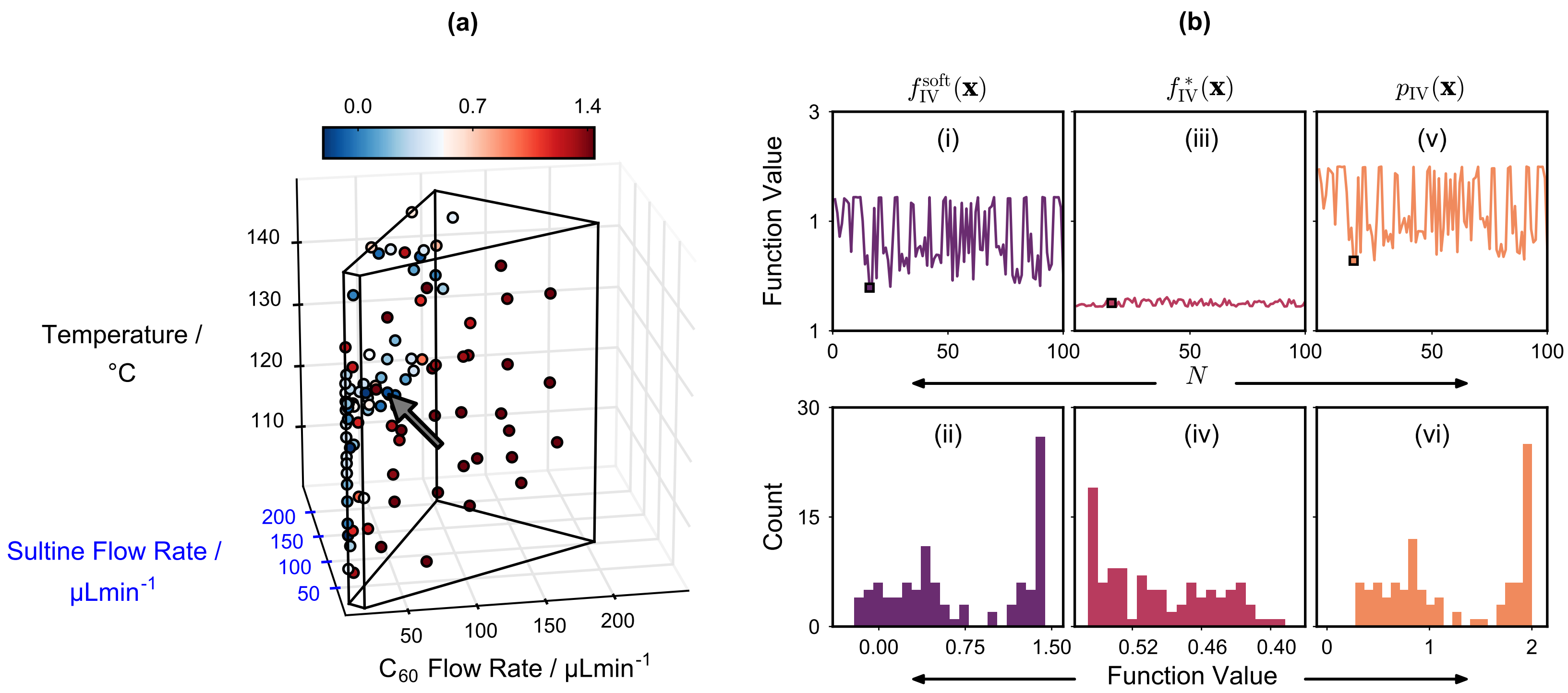

Figure S4: (a) Scatter plot for Run IV, showing the influence of the sultine flow rate, the C_{60} flow rate and the temperature on the merit value f_{IV}^{soft}. The location of each data point indicates the reaction conditions used, while the colour denotes the corresponding merit value. For ease of interpretation, points with merit values above the median merit value (0.522) have been coloured red, while those with merit values below the median value have been coloured blue; points with merit values close to the median value appear as white. The black cage defines the flow rate and temperature constraints. The arrows indicate the locations of the best point (i.e. the point with the lowest merit value) for the two runs. (b) Merit values for Run IV expressed as a time series (i) and a histogram (ii). Also shown are the time series and histograms for $f_{IV}^{*}(x)$ (iii, iv) and $p_{IV}(x)$ (v, vi). The square-shaped markers in the time-series plots correspond to the best point. (Note, for this run, no feasible point was found.)

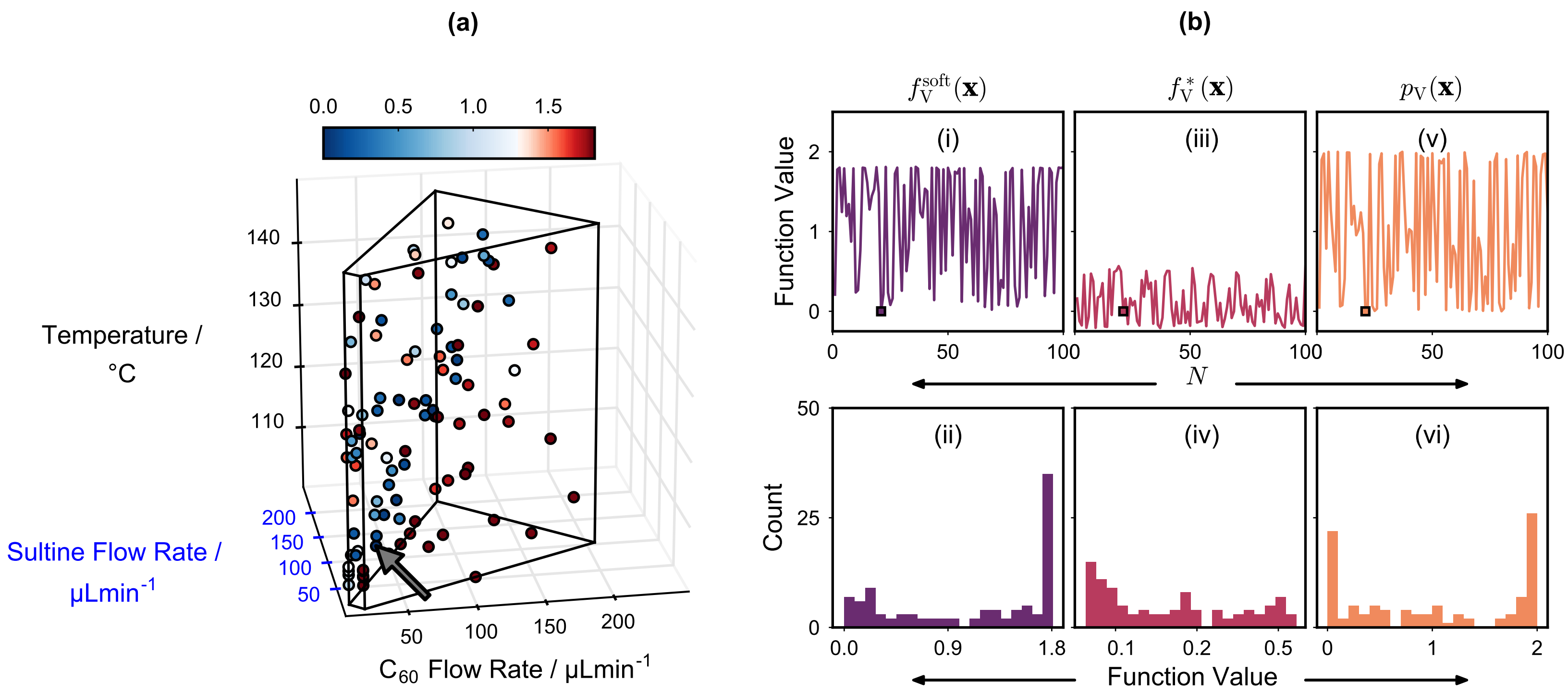

Figure S5: (a) Scatter plot for Run V, showing the influence of the sultine flow rate, the C_{60} flow rate and the temperature on the merit value $f_{\mathrm{V}}^{\mathrm{soft}}$. The location of each data point indicates the reaction conditions used, while the colour denotes the corresponding merit value. For ease of interpretation, points with merit values above the median merit value (1.326) have been coloured red, while those with merit values below the median value have been coloured blue; points with merit values close to the median value appear as white. The black cage defines the flow rate and temperature constraints. The arrows indicate the locations of the best point (i.e. the point with the lowest merit value) for the two runs. (b) Merit values for Run V expressed as a time series (i) and a histogram (ii). Also shown are the time series and histograms for $f_{\mathrm{V}}^{*}(x)$ (iii, iv) and $p_{\mathrm{V}}(x)$ (v, vi). The square-shaped markers in the time-series plots correspond to the best point, which was also the best feasible point.

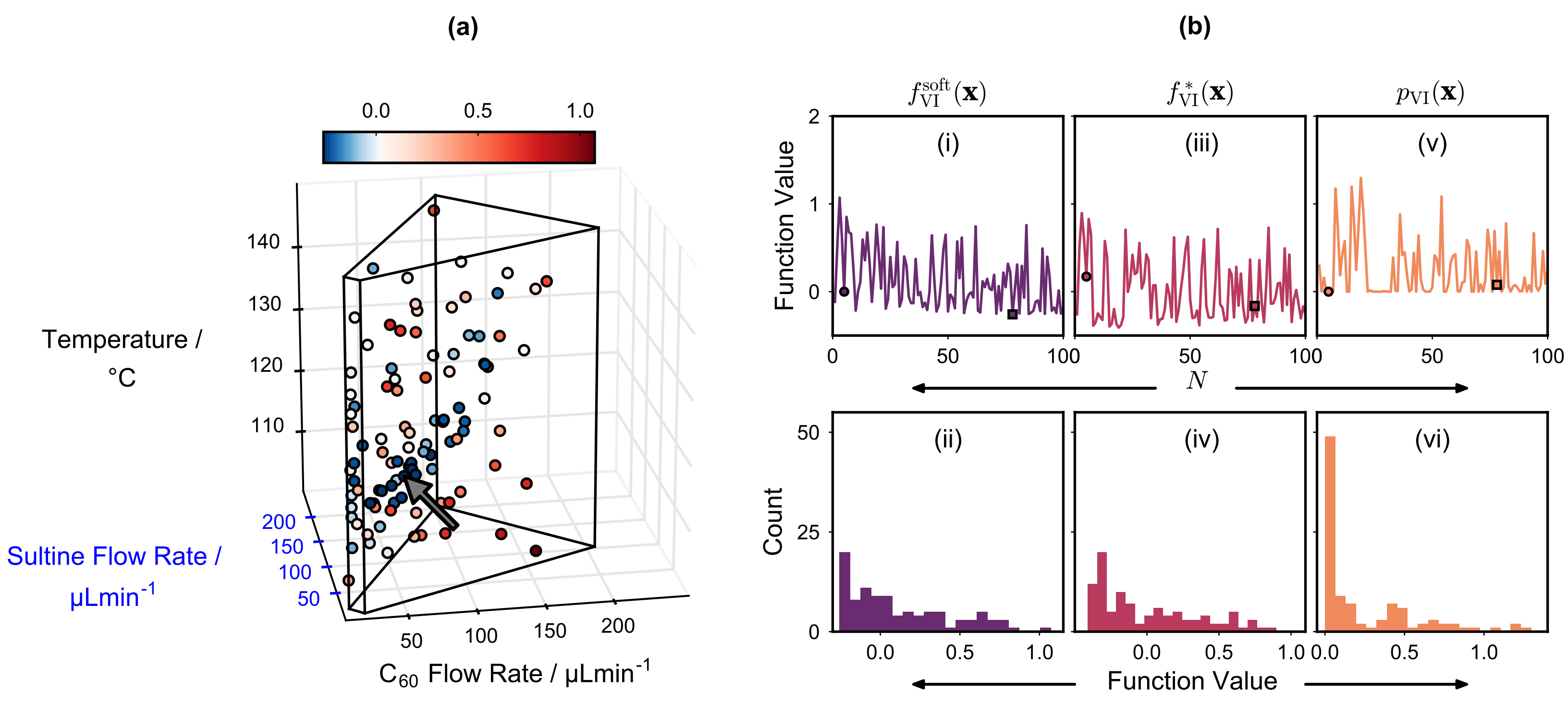

Figure S6: (a) Scatter plot for Run VI, showing the influence of the sultine flow rate, the C_{60} flow rate and the temperature on the merit value $f_{\mathrm{VI}}^{\mathrm{soft}}$. The location of each data point indicates the reaction conditions used, while the colour denotes the corresponding merit value. For ease of interpretation, points with merit values above the median merit value (0.0219) have been coloured red, while those with merit values below the median value have been coloured blue; points with merit values close to the median value appear as white. The black cage defines the flow rate and temperature constraints. The arrows indicate the locations of the best point (i.e. the point with the lowest merit value) for the two runs. (b) Merit values for Run VI expressed as a time series (i) and a histogram (ii). Also shown are the time series and histograms for $f_{\mathrm{VI}}^{*}(x)$ (iii, iv) and $p_{\mathrm{VI}}(x)$ (v, vi). The square and circle-shaped markers in the time-series plots correspond to the best point and the best feasible point, respectively.

Run		I	II	III	IV	V	VI
Optimisation Settings	**Objective Function**	$[X_3]$	$[X_3]$	$[X_3]$	$[X_3]$	$[X_3]$	$[X_1]$ / $[X_2]$
	Constraint(s)		$[X_1] + [X_2] > 0.9$	[X1] + [X2] > 0.9 $[X_1] / [X_2] > 4$	[X1] + [X2] > 0.9 $[X_1] / [X_2] < 0.5$	[X1] + [X2] > 0.9 $[X_1] / [X_2] < 1.5$	[X1] + [X2] > 0.9
	ΔF		0.3	0.3	0.3	0.3	0.3
	f_0		0.0128	0.0104	0.8546	0.0849	1.3931
	Δ		0.0434	0.0254	0.6753	0.3005	1.16849
	First Feasible Point		1	3		21	2
Best Point	**N**	53	33	26	16	21	78
	$f^*(x)$		-0.166	-0.206	-0.492	0.000	-0.324
	$p(x)$		0.001	0.024	0.280	0.000	0.078
	Merit Value (f or f^{soft})	0.001	-0.166	-0.182	-0.212	0.000	-0.258
	Reaction Time / min	3.0	5.1	4.8	5.8	12.1	7.3
	$[S] / [C_{60}]$	1.88	3.73	1.50	6.00	3.96	3.65
	Temperature / °C	105	104	119	125	107	116
	$[X_0]$	0.318	0.101	0.129	0.001	0.007	0.004
	$[X_1]$	0.649	0.784	0.776	0.288	0.535	0.382
	$[X_2]$	0.033	0.110	0.091	0.510	0.373	0.458
	$[X_3]$	0.001	0.004	0.004	0.201	0.085	0.157
Best Feasible Point	**N**		92	99		21	5
	$f^*(x)$		-0.1228	-0.0394		0	0
	$p(x)$		0	0		0	0
	Merit Value (f or f^{soft})		-0.1228	-0.0394		0	0
	Reaction Time / min		22.1	6.2		12.1	6.7
	$[S] / [C_{60}]$		1.73	1.69		3.96	5.69
	Temperature / °C		103	119		107	102
	$[X_0]$		0.089	0.085		0.007	0.007
	$[X_1]$		0.786	0.778		0.535	0.527
	$[X_2]$		0.119	0.128		0.374	0.378
	$[X_3]$		0.007	0.009		0.085	0.087

Table S1: Pertinent characteristics for each of the six optimisation runs.

Some comments on the construction and properties of the merit function used for constrained optimisation

For the work described here constraints were applied using an analytical technique due to Huyer and Neumaier[1] which, building on earlier work by Dalwig et al.[2], involves the transformation of an unconstrained objective function $f(\mathbf{x})$ (for the lead property) into a new bounded function $f^*(\mathbf{x})$, onto which an easily constructed bounded constraint function $p(\mathbf{x})$ (for the other properties) may subsequently be appended:

$$f^*(\mathbf{x}) = \frac{f(\mathbf{x}) - f_0}{\Delta + |f(\mathbf{x}) - f_0|} \qquad \text{(S1)}$$

where f_0 is the value of the original objective function at a previously determined feasible point $\mathbf{x}_0$ (i.e. at a point that satisfies all constraints) and Δ is a fixed positive number. There is some flexibility in the choice of Δ, although it is convenient to set it equal to the median value of $|f(\mathbf{x}) - f_0|$ obtained from a small set of data points distributed throughout the parameter space.

The constraint function $p(\mathbf{x})$ is a compound penalty function that assumes a positive value – and so increases the value of the objective function – whenever one or more of the constraints is violated. Specifically, the constraint function takes the form:

$$p(\mathbf{x}) = 2\left(\frac{\sum_i p_i^2(\mathbf{x})}{1 + \sum_i p_i^2(\mathbf{x})}\right) \qquad \text{(S2)}$$

where the index i runs over all constrained properties and $p_i(\mathbf{x})$ denotes the corresponding penalty function. For each constrained property a window $[F_i^{\text{Lower}}, F_i^{\text{Upper}}]$ is specified that denotes the preferred lower and upper limits for the property, together with a positive variable ΔF_i that denotes the maximum allowed digression away from these limits. $p_i(\mathbf{x})$ is assigned a value $[F_i^{\text{Lower}} - F_i(\mathbf{x})]$ / ΔF_i, zero, or $[F_i(\mathbf{x}) - F_i^{\text{Upper}}]$ / ΔF_i according to whether $F_i(\mathbf{x})$ lies below, within, or above the bounds of the window. In other words, whenever a constrained property lies outside the preferred range, a positive penalty is applied of a size that is proportional to the digression from the nearest limit.

From (S1), $f^*(\mathbf{x})$ lies in the range [-1, 1], with negative values corresponding to points with $f(\mathbf{x}) < f_0$. From (S2), $p(\mathbf{x})$ lies in the range [0, 2], with a value of zero signifying feasible points for which all constraints are fully satisfied, positive values in the range [0, 1] signifying semi-feasible points for which no constraint is violated by more than its corresponding sigma value, and positive values greater than one signifying infeasible points for which at least one constraint is violated by more than its corresponding sigma value. Combining the two functions, we obtain a new objective function $f^{\text{soft}}(\mathbf{x})$ spanning the range [-1, 3] that takes into account both the value of the lead property and the constraints:

$$f^{\text{soft}}(\mathbf{x}) = f^*(\mathbf{x}) + p(\mathbf{x}) \qquad \text{(S3)}$$

The constrained optimisation is carried out by searching for the location $\hat{\mathbf{x}}$ of the global minimum of $f^{\text{soft}}(\mathbf{x})$ - a straightforward procedure since, providing all properties vary smoothly with $\mathbf{x}$, Eq. (S3) yields a continuously differentiable merit function that may be readily minimised using standard optimisers.

From the above discussion, all feasible points (of which there is assumed to be at least one at $\mathbf{x}_0$) have f^{soft}-values less than or equal to one, while all infeasible points have f^{soft}-values greater than zero. $\hat{\mathbf{x}}$ cannot therefore be an infeasible point since, by construction of Eq. (S1), at least one feasible point has a lower f^{soft}-value of zero (and other feasible points may have negative

f^{soft}-values); moreover $\hat{\mathbf{x}}$ can only be a semi-feasible point if the reduction in $f^*(\mathbf{x})$ achieved by straying outside the feasible zone exceeds the penalty $p(\hat{\mathbf{x}})$ due to the transgression. Given the monotonic relationship between $f^*(\mathbf{x})$ and $f(\mathbf{x})$, it follows that $\hat{\mathbf{x}}$ will only be a semi-feasible point if $f(\hat{\mathbf{x}})$ is smaller than the lowest known value of $f(\mathbf{x})$ inside the feasible zone – a sensible basis for violating the constraint(s).

The individual penalty functions in (S2) merely require one to specify the preferred bounds for each property, together with a ΔF value defining the maximum permissible digression from these bounds – easily supplied quantities in many cases.* Hence the complete penalty function can be readily constructed on the basis of easily acquired physical information. This contrasts with conventional penalty functions where the penalty parameters (used to control the extent to which the constraints may be violated) are purely mathematical parameters with no direct physical interpretation, meaning their values cannot be set *a-priori* and must instead be decided through extensive experimentation[3].

The only other information required to carry out the constrained optimisation is a feasible point, which may be found by first minimising $p(\mathbf{x})$ until a value of zero is obtained. Alternatively, to avoid the need for a separate optimisation, measurements may be carried out at a small set of initial data points that are evenly distributed throughout the parameter space (the same ones used to determine Δ). If any of these points are found to be feasible, f_0 is set equal to the objective function value of the best feasible point in the set; otherwise it is set equal to $2f(\mathbf{x}) - f_{\min}$ as a temporary proxy for a feasible point, and the optimisation proceeds until a genuine feasible point is found. At this point f_0, Δ and all the f^{soft} values so far obtained are recalculated, before resuming the optimisation.† Either way, the optimisation routine requires little prior knowledge of the system and, by using only easily supplied values for the constraints, avoids the usual need for laborious tuning of merit function parameters.

References

1. W. Huyer and A. Neumaier, *ACM Trans. Mat. Soft.*, 2008, **35**, 1-25.

2. S. Dallwig, A. Neumaier and H. Schichl, in *Developments in global optimization*, ed. I. M. Bomze, T. Csendes, R. Horst and P. M. Pardalos, Springer US, 1997, pp. 19-36.

3. Ö. Yeniay, *Math. Comput. Appl.*, 2005, **10**(1), 45-56.

*In cases where achievable bounds are not known in advance, they may typically be determined by carrying out an initial optimisation run using sensibly guessed constraints and using the data generated during the course of that run to determine appropriate constraints for future runs.

†The heuristic of using $2f_{\max} - f_{\min}$ in place of a feasible point has been found to work for a broad range of numerical problems.