Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Material

Jennifer M. Elward and B. Christopher Rinderspacher*

E-mail: berend.c.rinderspacher.civ@mail.mil

*To whom correspondence should be addressed

Enumeration Convention

The compound $\{x_i = X_i\}_{i=1}^{8}$ is assigned the number,

$$n(x) = \sum_{i=1}^{M} X_i \prod_{j=0}^{i-1} N_j. \tag{1}$$

Where N_j is the number of options on substitution site j, and $N_0 = 1$. The naming scheme was chosen to reflect a general base numbering system. The choice of general base was made so as to directly relate the framework substitution options to the enumerated compound. The numbering for each substituent is shown in 1.

Table 1: Substituent numbering for general base enumeration scheme.

Substituent	**Number**
H	0
F	1
Br	2
Cl	3

An explicit example of the enumeration scheme used in this work is given here. Structure 9004, in 1 is evaluated. X_1, X_4, X_6 and X_8 all contribute 0 due to their hydrogen substi-

Table 2: Substituent evaluation for structure 9004.

Site	**Substituent**	**Numerical Value**	**Sum Contribution**
X_1	H	0	0
X_2	Br	3	3×4
X_3	Cl	2	2×4^2
X_4	H	0	0
X_5	Br	3	3×4^4
X_6	H	0	0
X_7	Cl	2	2×4^6
X_8	H	0	0

tution. Site X_2 has numerical value 3 for it's bromine substitution. The value for X_2 is then multiplied by the number of substitution options for site 2, which is four in total giving a total contribution from site 2 of $3 \times (4options) = 12$. The process is repeated for the

remaining sites, each of which are shown in 2. Summing up all contributions from column 4 leads to an enumerated value of 9004.

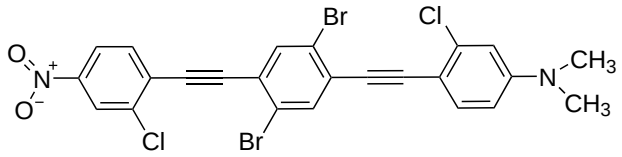


Figure 1: Enumerated structure 9004.

Final Structure Images

Below are 22 unique final structure images for each of the tested algorithms. The structures are given in terms of their respective starting points. Within each of those plots the chemical graphs for each of the structures are organized by structure number. The structures resulting from starting point zero are given in 2, starting point 4710 in 3, starting point 8389 in 4, starting point 41329 in 5 and finally starting point 41668 in 6.

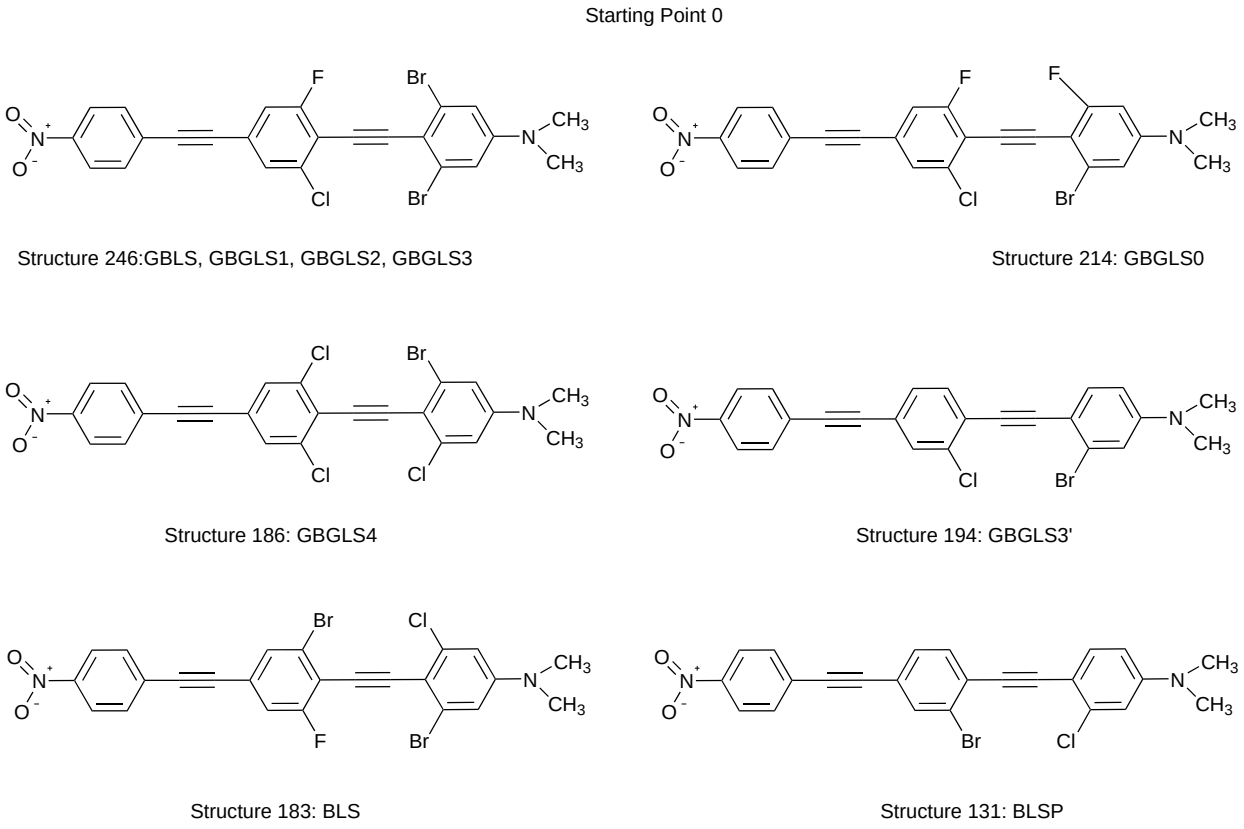
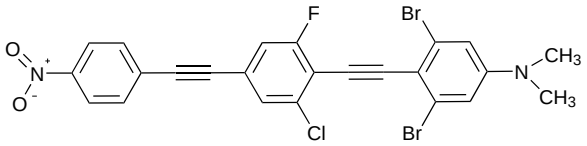
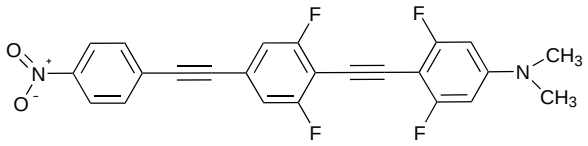


Figure 2: Collected final structures resulting from starting point 0, organized by structure number.

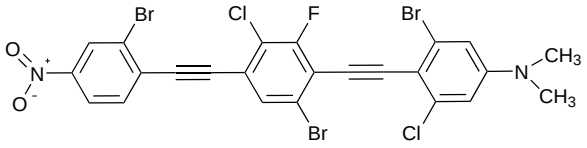
Starting Point 4710



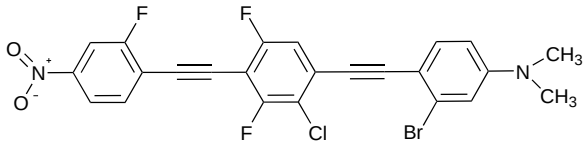
Structure 249: GBLS, GBGLS1, GBGLS2, GBGLS3, GBGLS4, GBGLS3'



Structure 85: GBGLS0



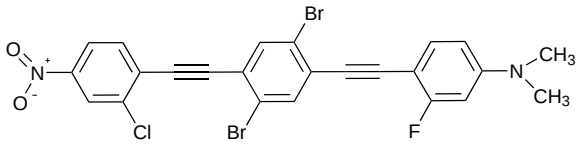
Structure 13037: BLS



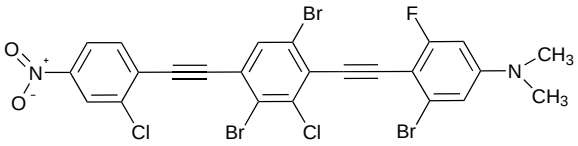
Structure 12821: BLSP

Figure 3: Collected final structures resulting from starting point 4710, organized by structure number.

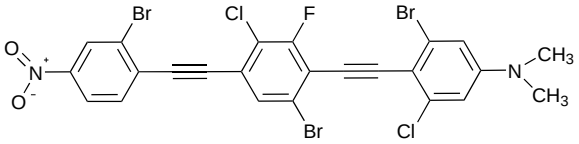
Starting Point 8389



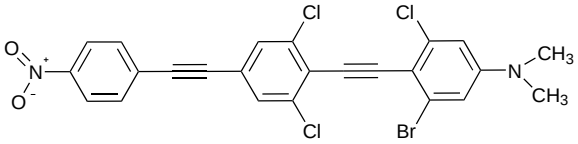
Structure 9080: GBLS, BLS



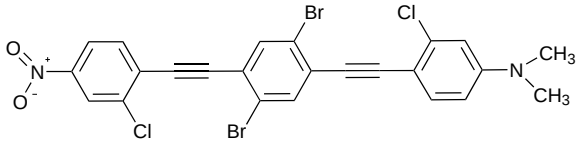
Structure 9182: GBGLS0



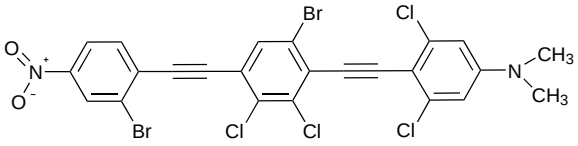
Structure 13037: GBGLS1, GBGLS2, GBGLS3



Structure 234: GBGLS4



Structure 9004: GBGLS3'



Structure 12974: BLSP

Figure 4: Collected final structures resulting from starting point 8389, organized by structure number.

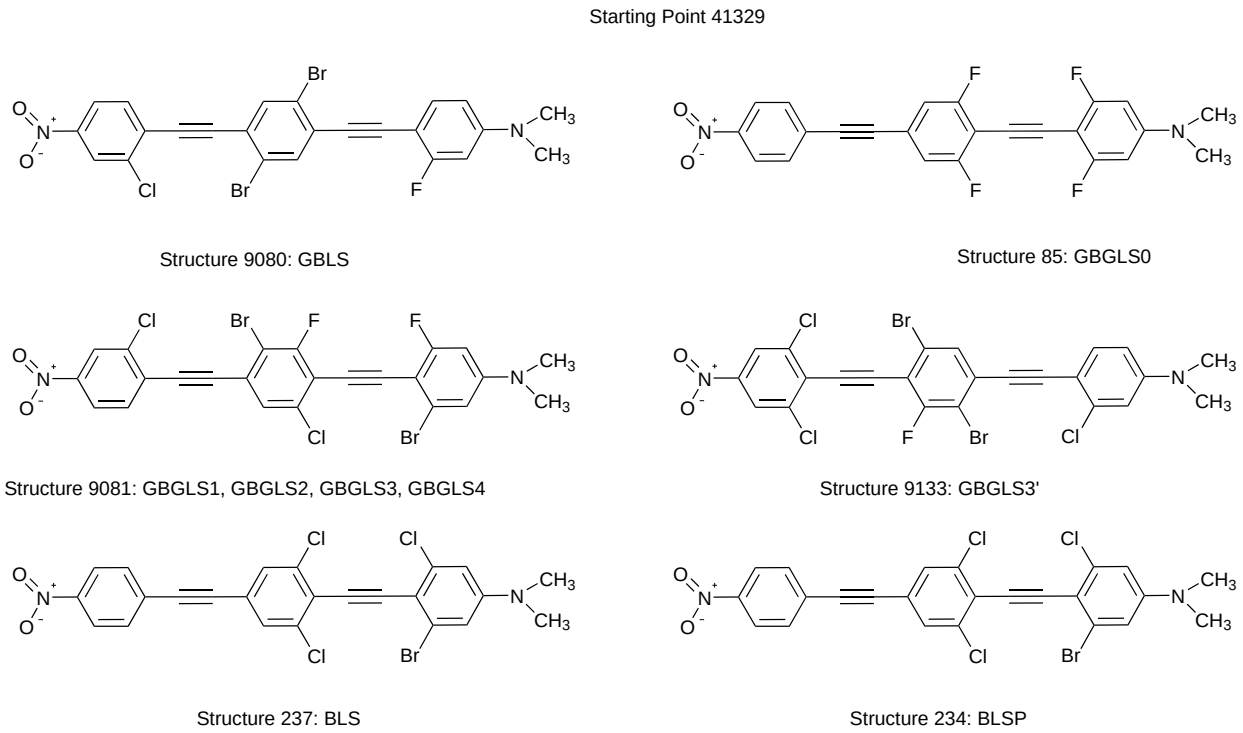
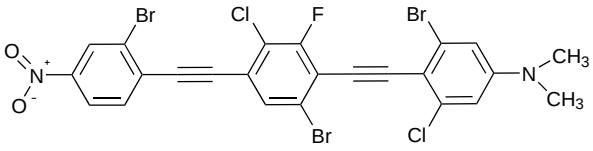
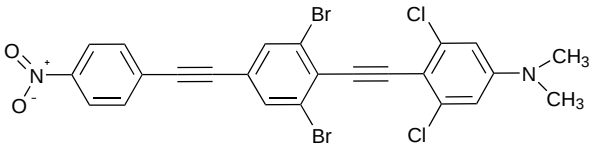


Figure 5: Collected final structures resulting from starting point 41329, organized by structure number.

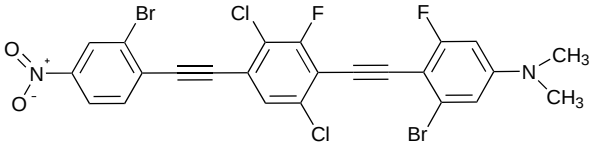
Starting Point 41668



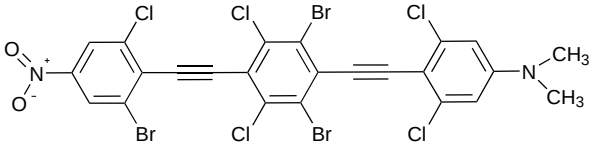
Structure 13037: GBLS, GBGLS1, GBGLS3



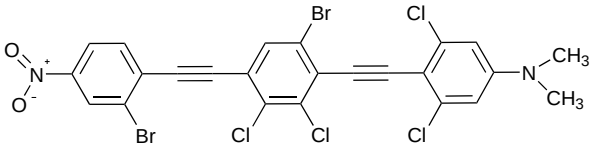
Structure 175: GBGLS0



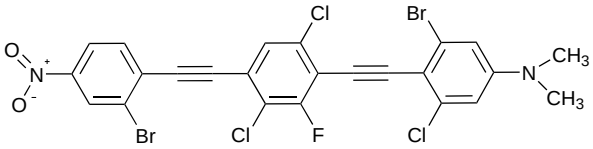
Structure 12921: GBGLS2, GBGLS4



Structure 45034: GBGLS3'



Structure 12974: BLS

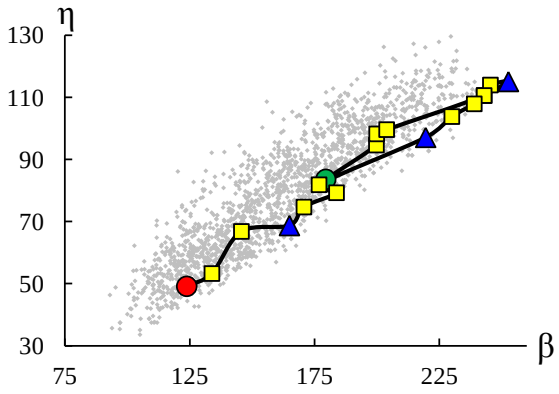


Structure 12985: BLSP

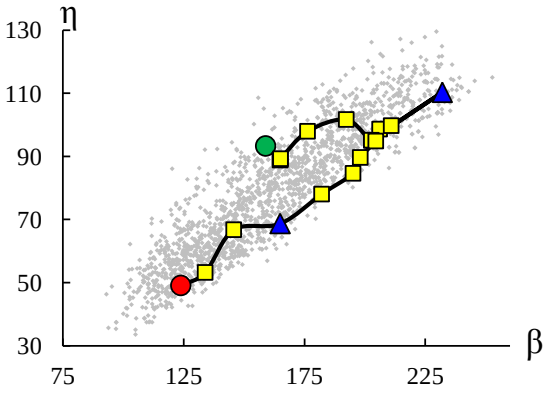
Figure 6: Collected final structures resulting from starting point 41668, organized by structure number.

Algorithmic Paths

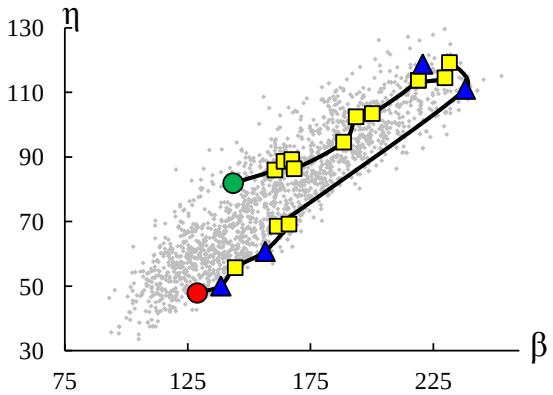
In this section, all algorithmic paths are given for each algorithm and starting point. For all of the plots given here, the hyperpolarizability (β) is given on the x-axis and the absorbance metric (η) is given on the y-axis. The green circles signify the start of the path and the red circle signifies the end of the path. The blue triangles denote λ parameter changes, and the yellow squares are optimization improvement points.



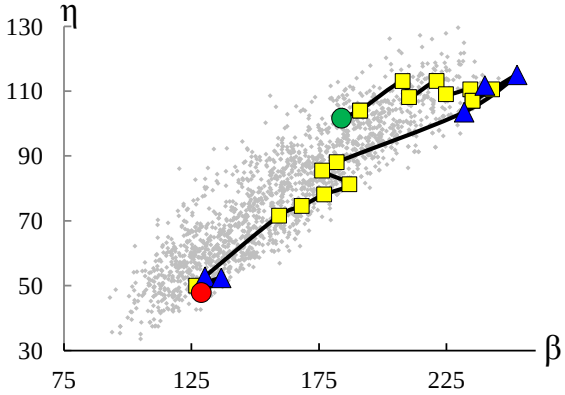
(a) Starting point 0, ending structure 246.



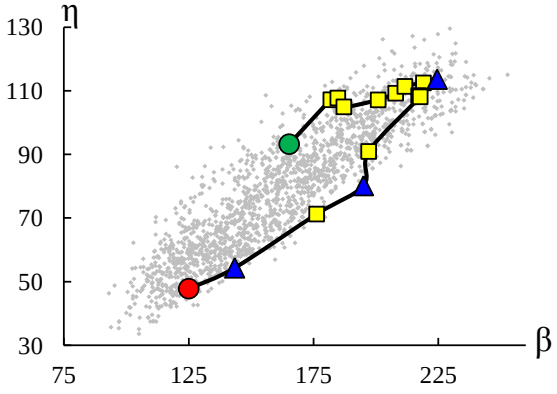
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 9080.

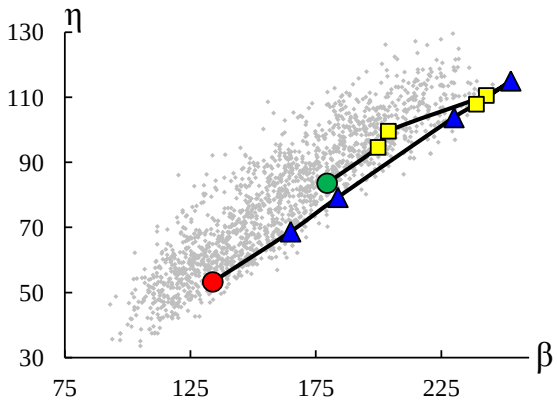


(d) Starting point 41329, ending structure 9080.

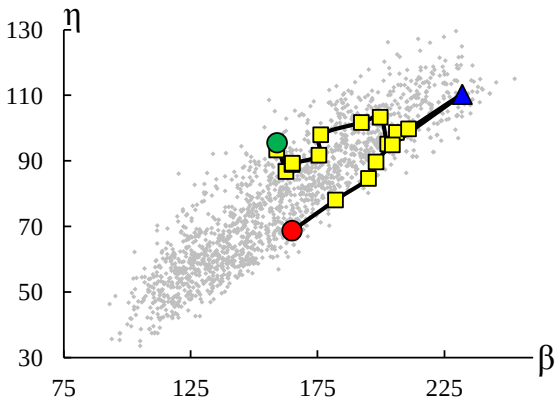


(e) Starting point 41668, ending structure 13037.

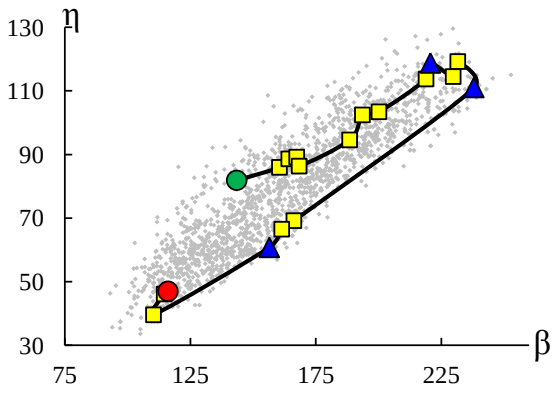
Figure 7: Paths resulting from GBLS algorithm.



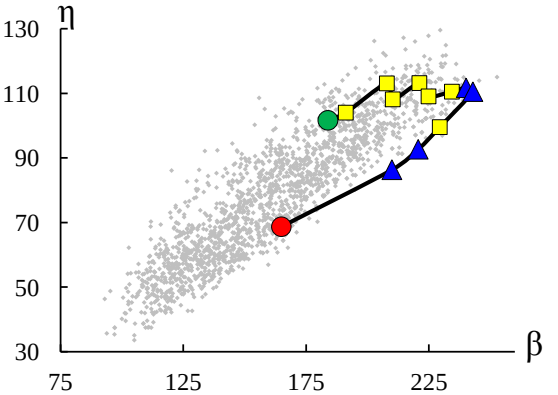
(a) Starting point 0, ending structure 214.



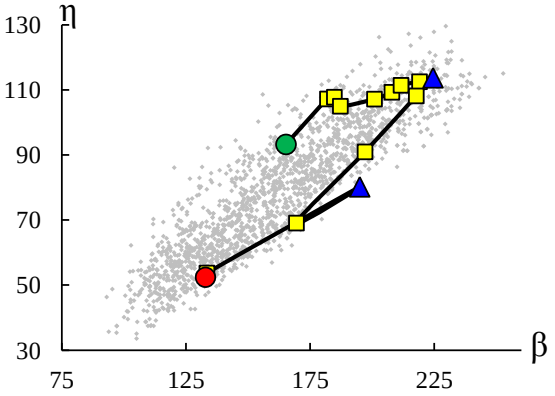
(b) Starting point 4710, ending structure 85.



(c) Starting point 8389, ending structure 9182.

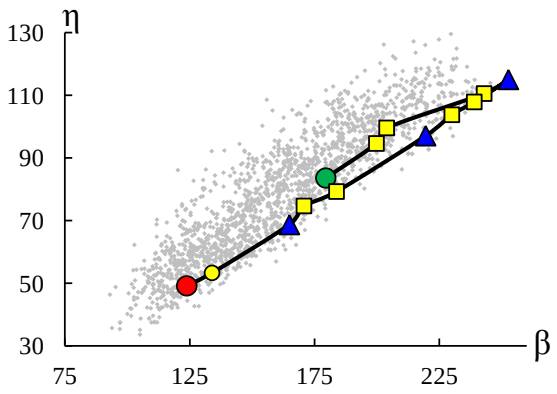


(d) Starting point 41329, ending structure 85.

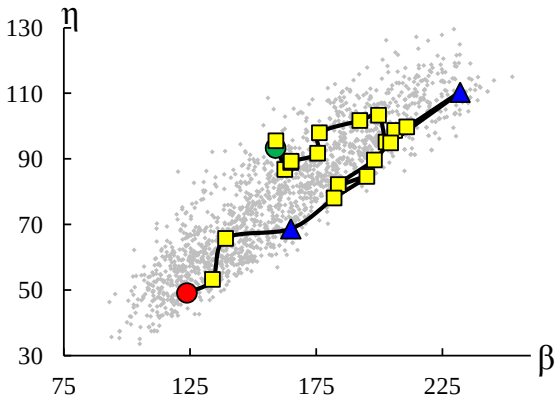


(e) Starting point 41668, ending structure 175.

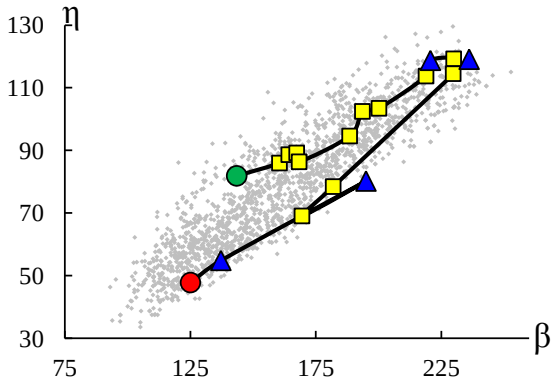
Figure 8: Paths resulting from GBGLS0 algorithm.



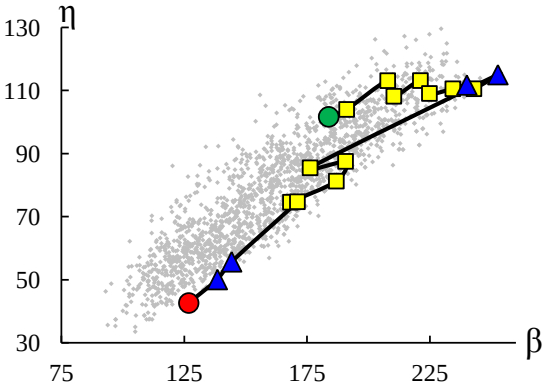
(a) Starting point 0, ending structure 246.



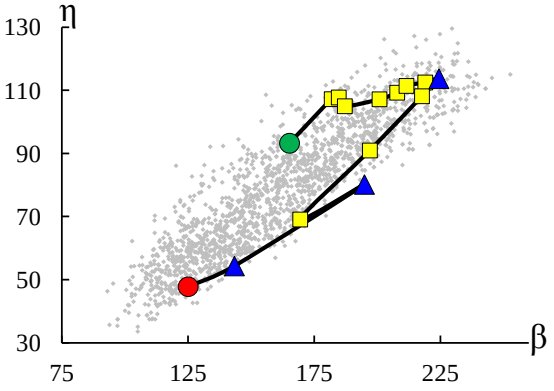
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 13037.

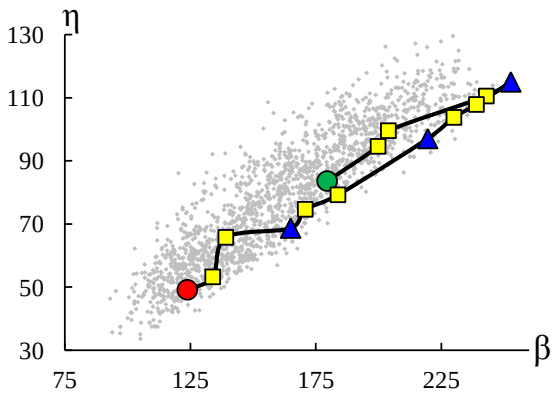


(d) Starting point 41329, ending structure 9081.

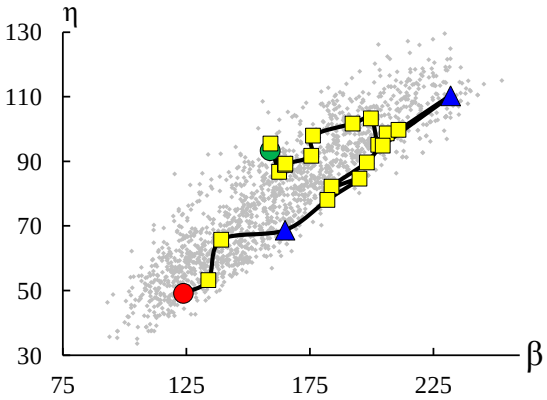


(e) Starting point 41668, ending structure 13037.

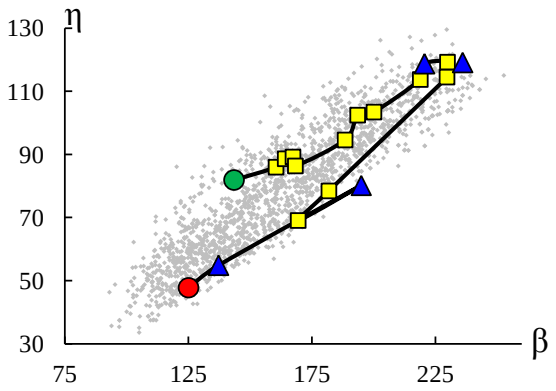
Figure 9: Paths resulting from GBGLS1 algorithm.



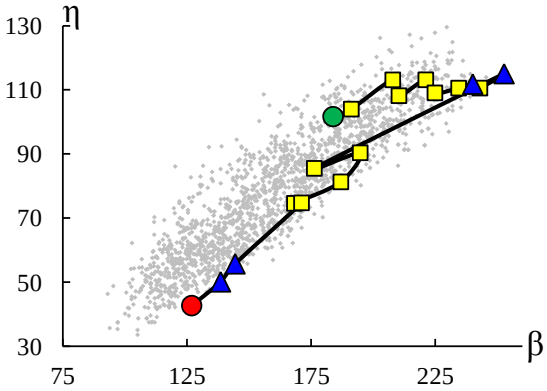
(a) Starting point 0, ending structure 246.



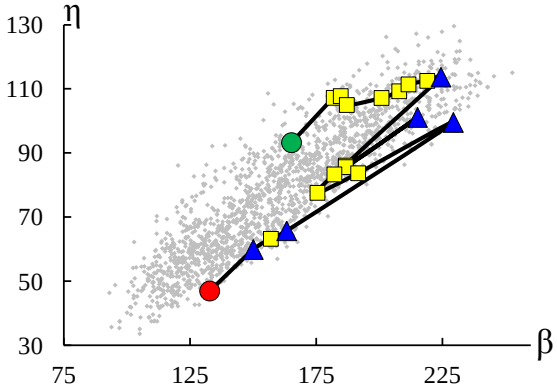
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 13037.

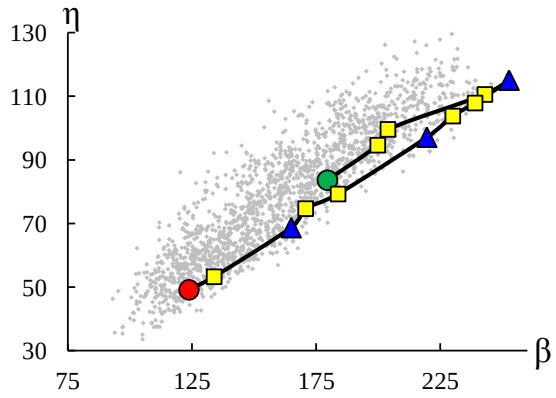


(d) Starting point 41329, ending structure 9081.

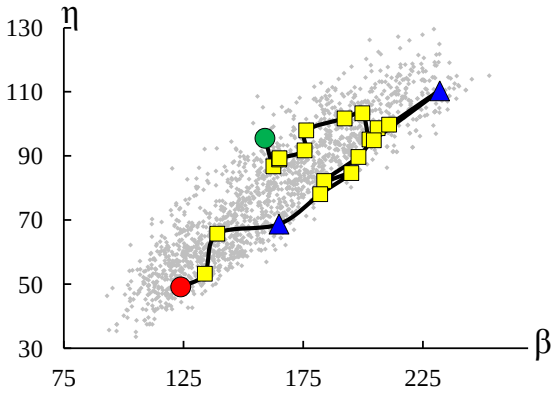


(e) Starting point 41668, ending structure 12921.

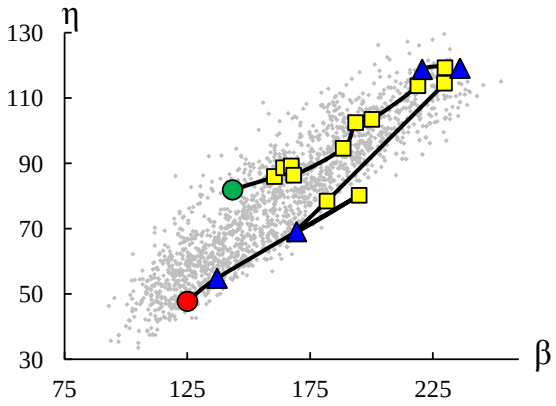
Figure 10: Paths resulting from GBGLS2 algorithm.



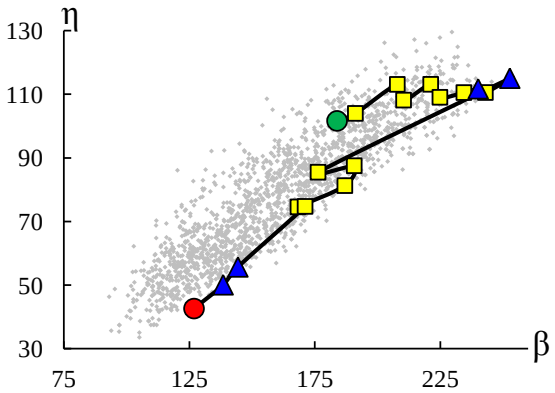
(a) Starting point 0, ending structure 246.



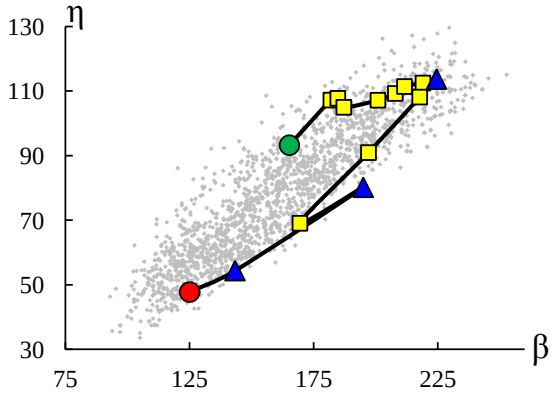
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 13037.

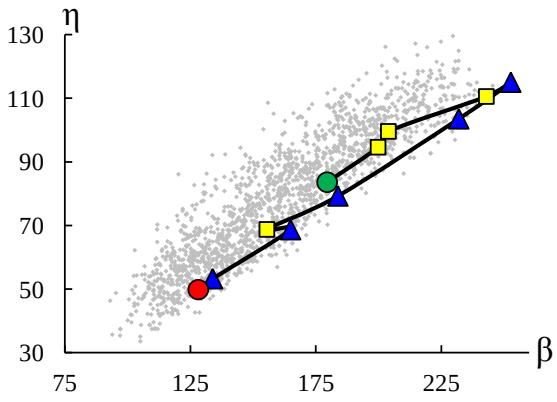


(d) Starting point 41329, ending structure 9081.

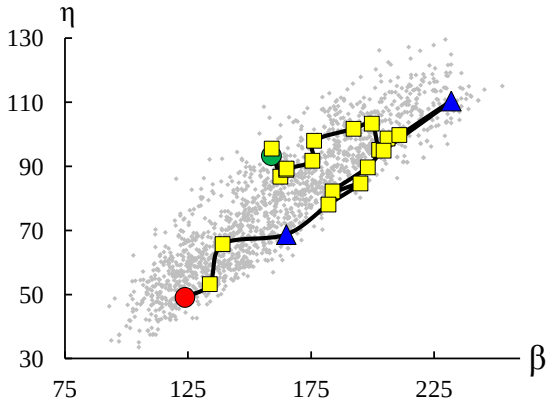


(e) Starting point 41668, ending structure 13037.

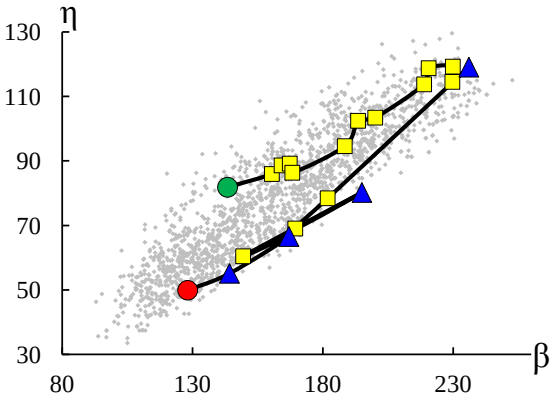
Figure 11: Paths resulting from GBGLS3 algorithm.



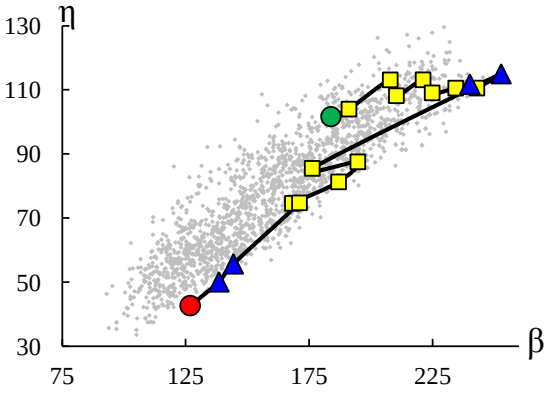
(a) Starting point 0, ending structure 186.



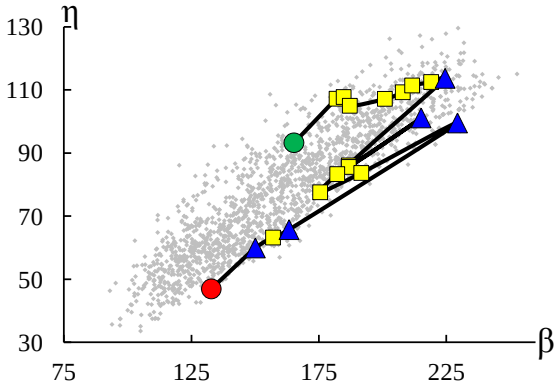
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 234.

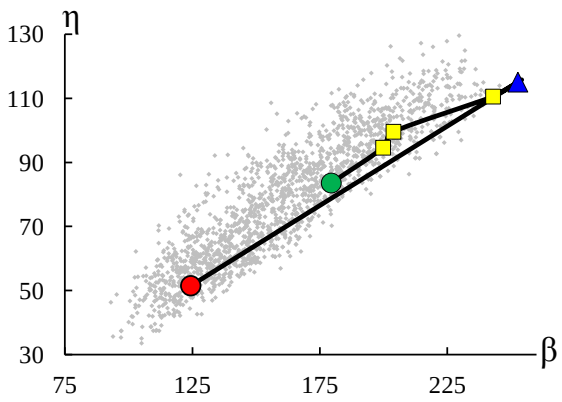


(d) Starting point 41329, ending structure 9081.

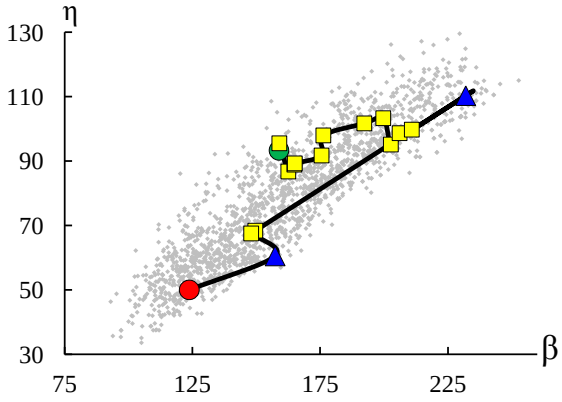


(e) Starting point 41668, ending structure 12921.

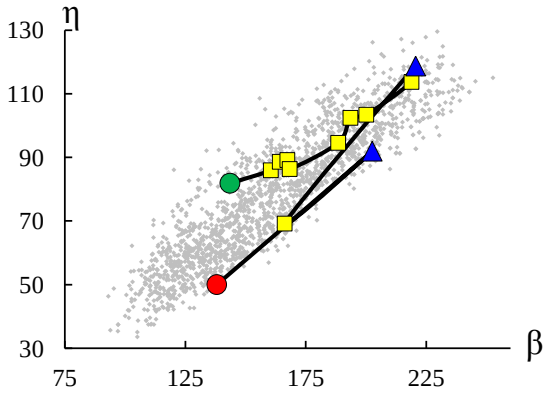
Figure 12: Paths resulting from GBGLS4 algorithm.



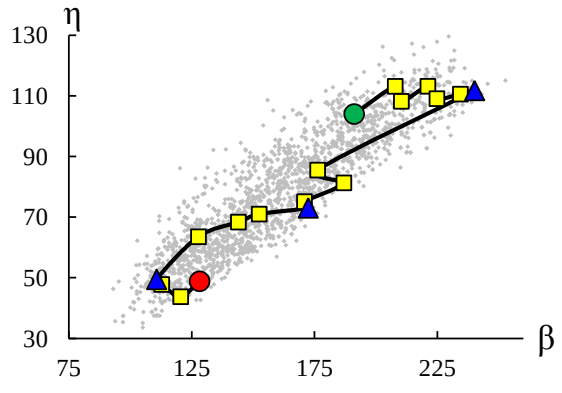
(a) Starting point 0, ending structure 194.



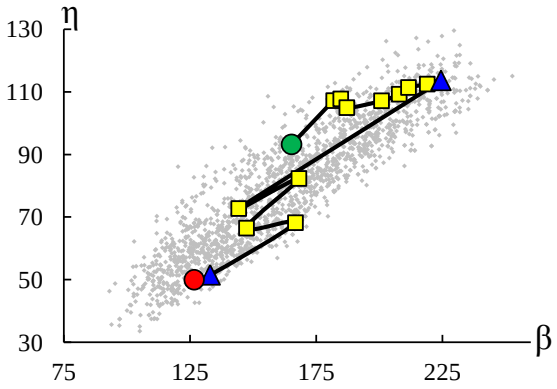
(b) Starting point 4710, ending structure 249.



(c) Starting point 8389, ending structure 9004.

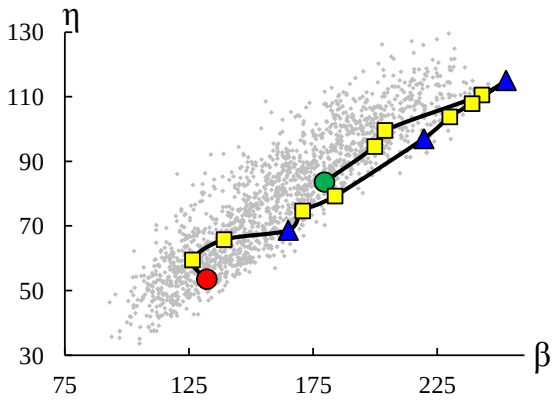


(d) Starting point 41329, ending structure 9133.

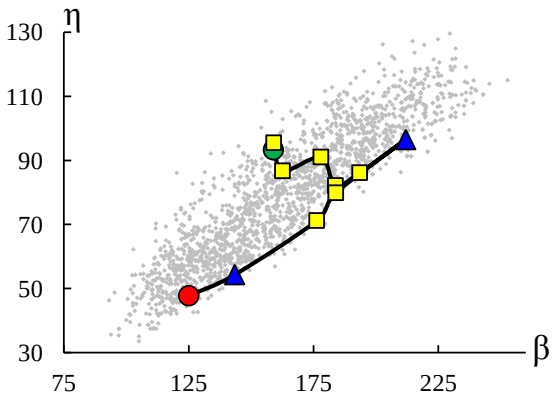


(e) Starting point 41668, ending structure 45034.

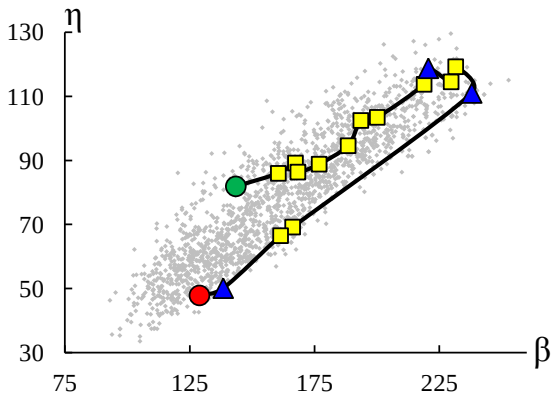
Figure 13: Paths resulting from GBGLS3′ algorithm.



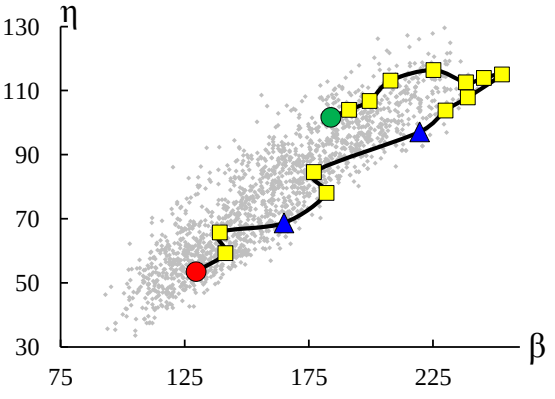
(a) Starting point 0, ending structure 183.



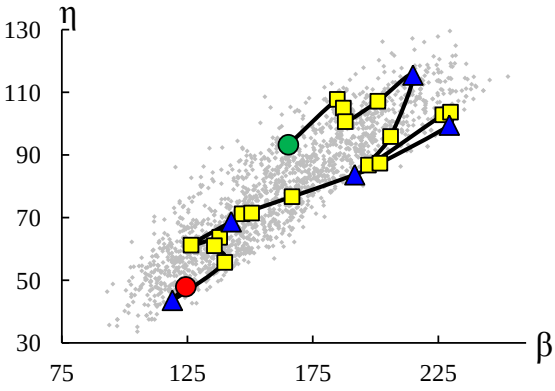
(b) Starting point 4710, ending structure 13037.



(c) Starting point 8389, ending structure 9080.

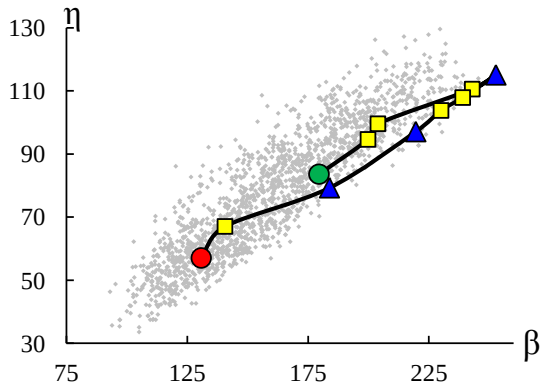


(d) Starting point 41329, ending structure 237.

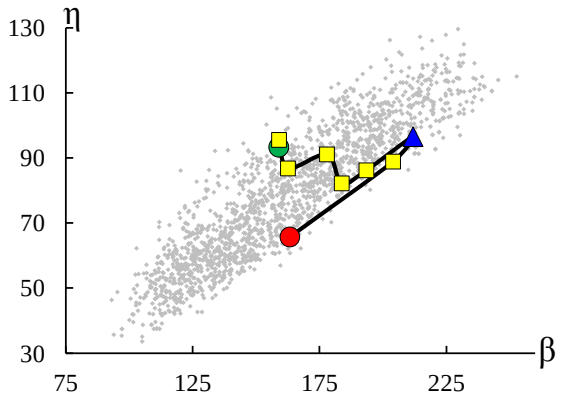


(e) Starting point 41668, ending structure 12974.

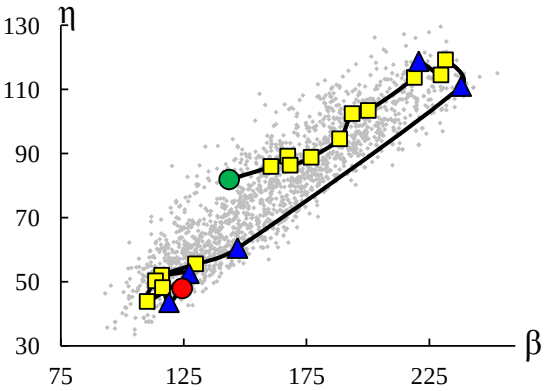
Figure 14: Paths resulting from BLS algorithm.



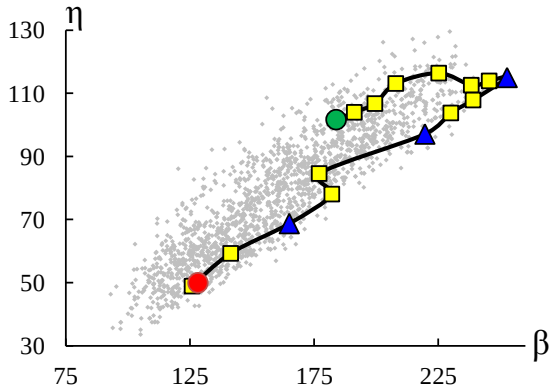
(a) Starting point 0, ending structure 131.



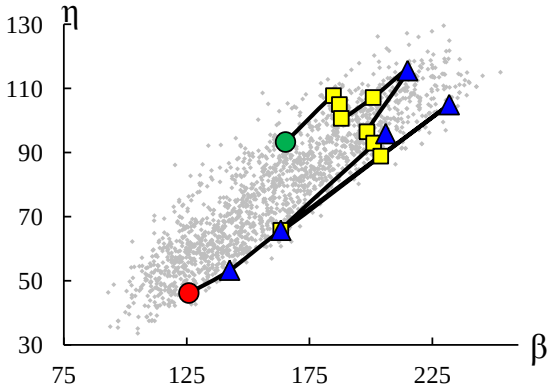
(b) Starting point 4710, ending structure 12821.



(c) Starting point 8389, ending structure 12974.



(d) Starting point 41329, ending structure 234.



(e) Starting point 41668, ending structure 12985.

Figure 15: Paths resulting from BLSP algorithm.

Lambda versus Lagrangian for all Algorithms

In this section, λ and Lagrangian comparison plots are given. To avoid repetition, a sampling of plots are given from each starting point including zero, 4710, 8389 and 41329. The plot for starting point 41668 is given in the main text. For clarity purposes, Lagrangian values given here have been scaled by the Δ factor in the amount shown in each plot's legend. The symbols used within the plots are as follows, the black circles denote GBGLS1 algorithm, the blue triangles denote GBGLS4 and GBGLS2 algorithm and they are noted as such, the red triangles denote GBGLS3′, the green squares denote GBGLS0, the light blue asterisks denote GBLS, the purple diamonds denote BLS algorithm, and the orange squares with black cross denote BLSP.

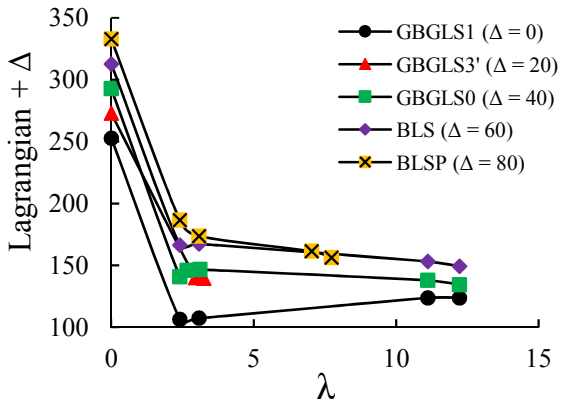


Figure 16: Lagrangian versus λ comparison for starting point zero.

The plot shown in 19 has an inset of λ versus Lagrangian for GBGLSX, where GBGLSX stands for algorithms GBGLS1 through GBGLS4. These particular algorithms were given as an inset due to the considerable difference in numerical value of λ.

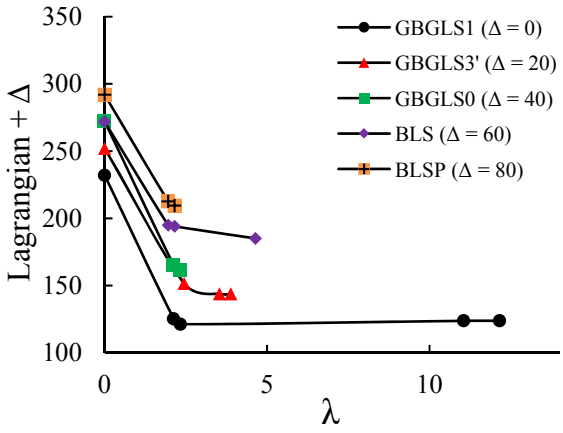


Figure 17: Lagrangian versus λ comparison for starting point 4710.

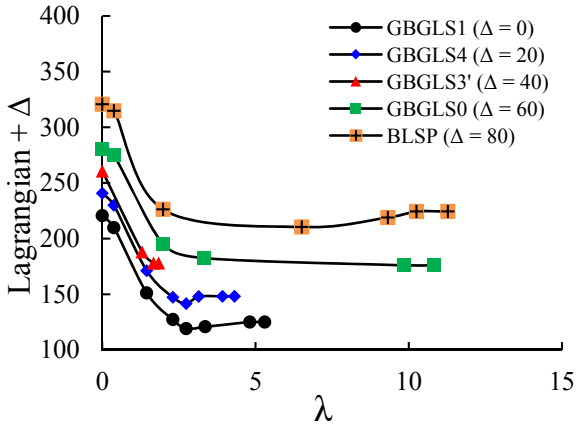


Figure 18: Lagrangian versus λ comparison for starting point 8389.

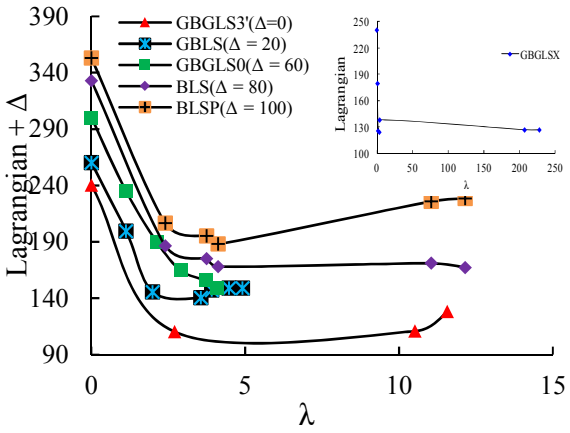
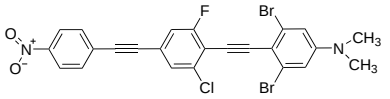
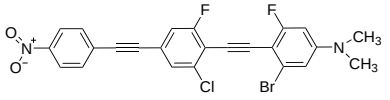
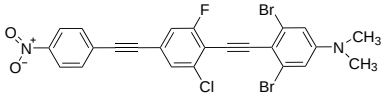
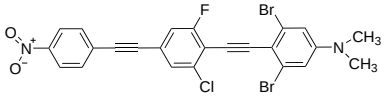
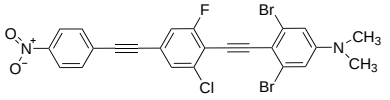
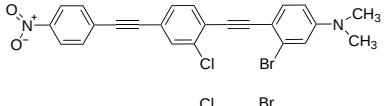
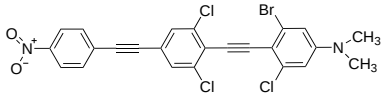
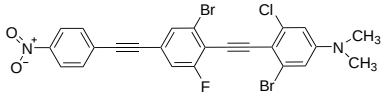


Figure 19: Lagrangian versus λ comparison for starting point 41329. Inset is plot of λ versus Lagrangian for GBGLSX=1-4.

Property Data Associated with Each Algorithm

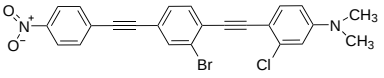
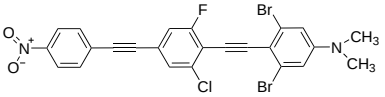
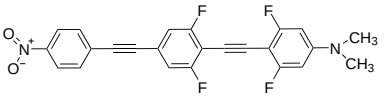
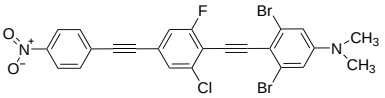
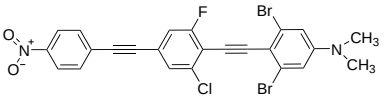
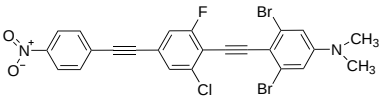
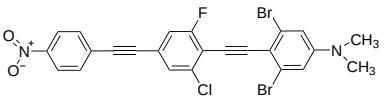
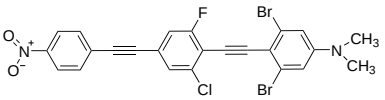
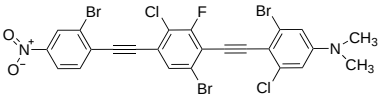
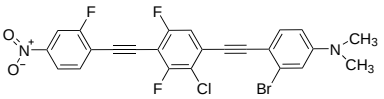
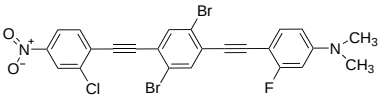
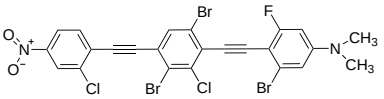
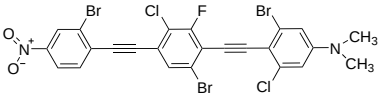
In 3 all data associated with each starting point is given. This includes, for each algorithm, the hyperpolarizability (β), the penalty for the run (π), the absorbance metric (η), the total number of molecules computed, the final structure number, and the chemical graph for the final structure.

Table 3: Constrained property (hyperpolarizability (β) given in units of 10^{-30}esu^{-1}), associated penalty (π) given in nm, absorbance metric (η), number of molecules computed (N. Mol.) and final structure for each algorithm studied. Five starting points have been examined.

Method	**Start**	β	π	η	**N. Mol.**	**Final**	**Chemical graph**
GBLS	0	123.8	0	49.1	101	246	
GBGLS0		133.8	3.2	53.2	84	214	
GBGLS1		123.8	0	49.1	81	246	
GBGLS2		123.8	0	49.1	80	246	
GBGLS3		123.8	0	49.1	79	246	
GBGLS3′		124.4	1.5	51.5	40	194	
GBGLS4		128.1	0	49.8	102	186	
BLS		132.1	3.5	53.5	69	183	

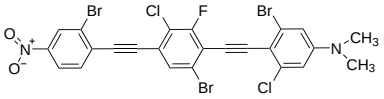
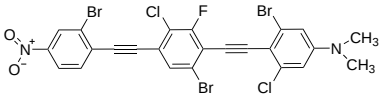
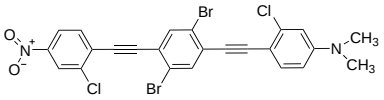
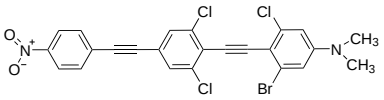
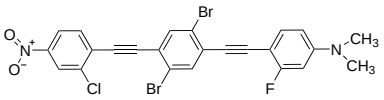
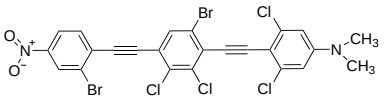
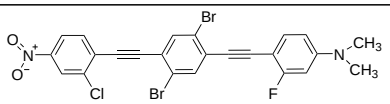
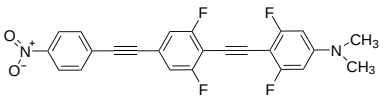
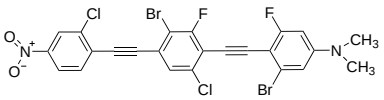
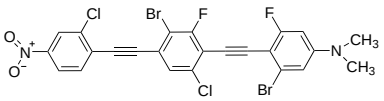
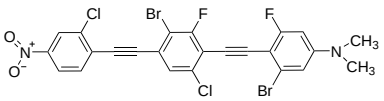
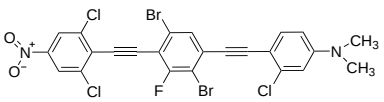
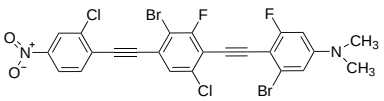
Continued on next page

Table 3 – *Continued from previous page*

Method	**Start**	β	π	η	**N. Mol.**	**Final**	**Chemical graph**
BLSP		130.8	7.0	57.0	79	131	
GBLS	4710	123.8	0	49.1	98	249	
GBGLS0		164.9	18.7	68.7	55	85	
GBGLS1		123.8	0	49.1	84	249	
GBGLS2		123.8	0	49.1	84	249	
GBGLS3		123.8	0	49.1	84	249	
GBGLS3′		123.8	0	49.1	67	249	
GBGLS4		123.8	0	49.1	84	249	
BLS		125.0	0	47.8	54	13037	
BLSP		163.3	15.7	65.7	48	12821	
GBLS	8389	128.9	0	47.9	119	9080	
GBGLS0		116.1	0	47.0	92	9182	
GBGLS1		125.0	0	47.8	124	13037	

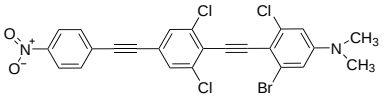
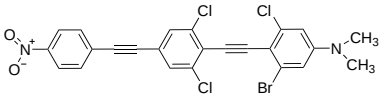
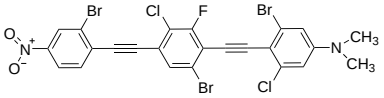
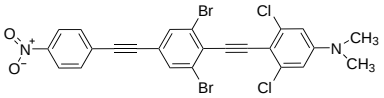
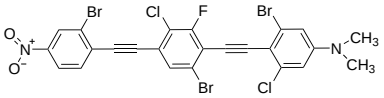
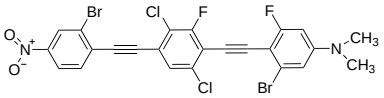
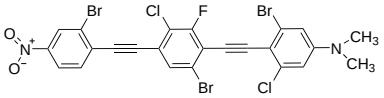
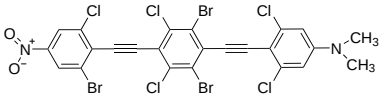
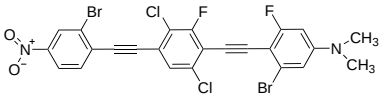
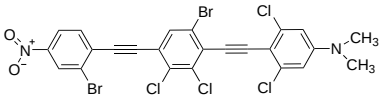
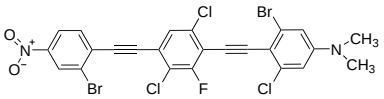
Continued on next page

Table 3 – *Continued from previous page*

Method	**Start**	β	π	η	**N. Mol.**	**Final**	**Chemical graph**
GBGLS2		125.0	0	47.8	124	13037	
GBGLS3		125.0	0	47.8	124	13037	
GBGLS3′		138.0	0	49.5	60	9004	
GBGLS4		128.2	0	49.9	130	234	
BLS		128.9	0	47.9	64	9080	
BLSP		124.3	0	47.9	111	12974	
GBLS	41329	128.9	0	47.9	140	9080	
GBGLS0		164.9	18.7	68.7	86	85	
GBGLS1		126.9	0	42.6	113	9081	
GBGLS2		126.9	0	42.6	113	9081	
GBGLS3		126.9	0	42.6	113	9081	
GBGLS3′		128.2	0	48.8	87	9133	
GBGLS4		126.9	0	42.6	113	9081	

Continued on next page

Table 3 – *Continued from previous page*

Method	**Start**	β	π	η	**N. Mol.**	**Final**	**Chemical graph**
BLS		129.5	3.5	53.5	83	237	
BLSP		128.2	0	49.9	112	234	
GBLS	41668	125.0	0	47.8	103	13037	
GBGLS0		132.8	2.5	52.4	87	175	
GBGLS1		125.0	0	47.8	82	13037	
GBGLS2		132.8	0	46.9	109	12921	
GBGLS3		125.0	0	47.8	82	13037	
GBGLS3′		126.6	0	48.0	61	45034	
GBGLS4		132.8	0	46.9	110	12921	
BLS		124.3	0	47.9	88	12974	
BLSP		125.8	0	46.1	109	12985	

In 4 through 8 all data for final substitution statistics are given in terms of starting point. Each algorithm tested is given here and the data is organized by substitution site and the final substitutions are ranked in terms of highest probability, lowest probability and two

intermediates.

Table 4: Final substitution site statistics for GBGLS1-4 and GBGLS3′, starting point zero.

Algorithm	Site Number	High	Low	Intermediate(1)	Intermediate(2)
GBGLS1	1	Cl	H	Br	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Cl	F	Br	H
	6	Cl	H	Br	F
	7	Cl	Br	H	F
	8	Cl	F	H	Br
GBGLS2	1	Cl	H	Br	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Cl	F	Br	H
	6	Cl	H	F	Br
	7	Cl	Br	H	F
	8	H	Cl	Br	F
GBGLS3	1	Cl	H	Br	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Cl	F	Br	H
	6	Cl	H	Br	F
	7	Cl	Br	H	F
	8	Cl	F	H	Br
GBGLS3′	1	Cl	H	F	Br
	2	F	Br	Cl	H
	3	Cl	H	F	Br
	4	Cl	H	Br	F
	5	H	Cl	F	Br
	6	H	Cl	F	Br
	7	H	Br	F	Cl
	8	H	Cl	F	Br
GBGLS4	1	Br	H	Cl	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Cl	Br	H	F
	6	H	Cl	Br	F
	7	H	Br	F	Cl
	8	H	F	Cl	Br

Table 5: Final substitution site statistics for GBGLS1-4 and GBGLS3′, starting point 4710.

Algorithm	Site Number	High	Low	Intermediate(1)	Intermediate(2)
GBGLS1	1	F	Br	Cl	H
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	F	Cl	H
	5	H	Cl	Br	F
	6	Cl	F	Br	H
	7	H	F	Cl	Br
	8	F	H	Cl	Br
GBGLS2	1	F	Br	Cl	H
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	F	Cl	H
	5	H	Cl	Br	F
	6	Br	F	Cl	H
	7	H	F	Cl	Br
	8	F	H	Cl	Br
GBGLS3	1	F	Br	Cl	H
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	F	Cl	H
	5	H	Cl	Br	F
	6	Cl	F	Br	H
	7	H	F	Cl	Br
	8	F	H	Cl	Br
GBGLS3′	1	F	H	Cl	Br
	2	Br	F	H	Cl
	3	Br	H	F	Cl
	4	Br	H	F	Cl
	5	H	Br	Cl	F
	6	Cl	Br	H	F
	7	H	Br	Cl	F
	8	H	F	Cl	Br
GBGLS4	1	F	Br	Cl	H
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	F	Cl	H
	5	H	Cl	Br	F
	6	Br	F	Cl	H
	7	H	F	Cl	Br
	8	F	H	Cl	Br

Table 6: Final substitution site statistics for GBGLS1-4 and GBGLS3′, starting point 8389.

Algorithm	Site Number	High	Low	Intermediate (1)	Intermediate (2)
GBGLS1	1	Cl	H	F	Br
	2	F	H	Br	Cl
	3	Cl	F	Br	H
	4	Cl	H	Br	F
	5	Br	F	Cl	H
	6	H	F	Cl	Br
	7	F	Cl	H	Br
	8	H	F	Cl	Br
GBGLS2	1	Cl	H	F	Br
	2	F	H	Br	Cl
	3	Cl	F	Br	H
	4	Cl	H	Br	F
	5	Br	F	Cl	H
	6	H	Br	Cl	F
	7	F	Cl	H	Br
	8	H	F	Cl	Br
GBGLS3	1	Cl	H	F	Br
	2	F	H	Br	Cl
	3	Cl	F	Br	H
	4	Cl	H	Br	F
	5	Br	F	Cl	H
	6	H	F	Cl	Br
	7	F	Cl	H	Br
	8	H	F	Cl	Br
GBGLS3′	1	F	Br	Cl	H
	2	H	F	Br	Cl
	3	Cl	H	Br	F
	4	F	Br	H	Cl
	5	Cl	F	Br	H
	6	H	Cl	F	Br
	7	Cl	H	F	Br
	8	F	H	Br	Cl
GBGLS4	1	Cl	Br	F	H
	2	Cl	H	F	Br
	3	Cl	F	Br	H
	4	Cl	H	F	Br
	5	H	F	Cl	Br
	6	Cl	Br	H	F
	7	F	Cl	H	Br
	8	H	F	Br	Cl

Table 7: Final substitution site statistics for GBGLS1-4 and GBGLS3′, starting point 41329.

Algorithm	Site Number	High	Low	Intermediate(1)	Intermediate(2)
GBGLS1	1	Br	H	Cl	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Br	F	Cl	H
	6	H	F	Br	Cl
	7	Cl	F	H	Br
	8	H	F	Br	Cl
GBGLS2	1	Br	H	Cl	F
	2	Cl	H	Br	F
	3	Cl	H	Br	F
	4	Cl	H	Br	F
	5	Br	F	Cl	H
	6	H	Cl	Br	F
	7	Cl	F	H	Br
	8	H	F	Br	Cl
GBGLS3	1	Br	H	Cl	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Br	F	Cl	H
	6	H	F	Br	Cl
	7	Cl	F	H	Br
	8	H	F	Br	Cl
GBGLS3′	1	Cl	H	Br	F
	2	Cl	H	Br	F
	3	Cl	H	Br	F
	4	Cl	F	Br	H
	5	Cl	F	Br	H
	6	H	F	Br	Cl
	7	Cl	F	Br	H
	8	H	F	Br	Cl
GBGLS4	1	Br	H	Cl	F
	2	Cl	H	Br	F
	3	Br	H	Cl	F
	4	Br	H	Cl	F
	5	Br	F	Cl	H
	6	H	Cl	Br	F
	7	Cl	F	H	Br
	8	H	F	Br	Cl

Table 8: Final substitution site statistics for GBGLS1-4 and GBGLS3′, starting point 41668.

Algorithm	Site Number	High	Low	Intermediate(1)	Intermediate(2)
GBGLS1	1	F	Br	Cl	H
	2	Br	F	Cl	H
	3	Cl	H	Br	F
	4	Cl	H	Br	F
	5	Cl	F	H	Br
	6	Cl	Br	H	F
	7	F	Cl	Br	H
	8	H	Cl	Br	F
GBGLS2	1	F	Br	Cl	H
	2	Cl	Br	F	H
	3	Br	F	Cl	H
	4	Cl	Br	F	H
	5	Cl	F	Br	H
	6	Br	Cl	H	F
	7	Br	F	H	Cl
	8	H	Cl	Br	F
GBGLS3	1	F	Br	Cl	H
	2	Br	F	Cl	H
	3	Cl	H	Br	F
	4	Cl	H	Br	F
	5	Br	F	H	Cl
	6	Cl	Br	H	F
	7	F	Cl	Br	H
	8	H	Cl	Br	F
GBGLS3′	1	Cl	H	Br	F
	2	Cl	Br	H	F
	3	Cl	H	Br	F
	4	Br	H	Cl	F
	5	Br	F	H	Cl
	6	Cl	H	Br	F
	7	H	Br	Cl	F
	8	H	F	Br	Cl
GBGLS4	1	F	Br	Cl	H
	2	Cl	Br	F	H
	3	Br	F	Cl	H
	4	Cl	Br	F	H
	5	Cl	F	Br	H
	6	Br	Cl	H	F
	7	Br	F	H	Cl
	8	H	Cl	Br	F