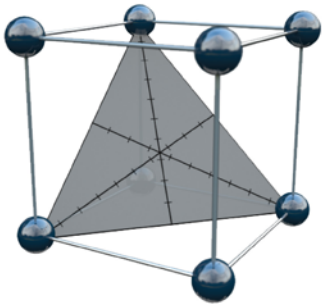


S-Matrix®

Fusion QbD Case Study – Capillary Electrophoresis Method Development

Exchanging HPLC for Capillary Electrophoresis in a Production Assay of Critical Product Yield

www.smatrix.com

BACKGROUND

This study was undertaken by a major international pharmaceutical company customer using the Fusion Process Development™ module of the Fusion QbD® Software Platform (S-Matrix Corporation, Eureka, CA, USA). The goal of the study was replacing an expensive high-performance liquid chromatography (HPLC) assay with a quick and substantially less expensive capillary electrophoresis (CE) assay. The assay was employed to monitor production yield in the manufacture of the API in a critical market product.

The study focused on four critical responses, each with acceptability limits for allowing the exchange of assays.

1. Mean Peak Area of the API
   Acceptability limit: ≥ 1,000 area counts.
2. Peak Area %RSD for the API
   Acceptability limit: ≤ 1.00%.
3. Assay Time (defined for this study as the Migration Time of the last peak).
   Acceptability limit: ≤ 14 minutes (current HPLC assay run time).
4. API Resolution
   Acceptability limit: ≥ 1.50.

EXPERIMENT DESIGN

The study included six CE assay parameters:

1. pH
2. Background Electrolyte ("BGE", millimoles)
3. Voltage (kV)
4. Pressure (millibars)
5. Surfactant (millimoles)
6. Visualizing Agent ("Vis. Agent", millimoles).

   A strong UV absorber. The API in this product is a very poor UV absorber, so an indirect UV assay was used. The Vis. Agent gives a high background UV absorbance. Reversing the CE voltage then gives a positive peak in the presence of the API (i.e., in the absence of Vis. Agent).

The study was carried out in two phases. In the first phase, three repeats each of seven "tuning runs" were carried out at various combinations of the six CE control factors to determine whether the required performance goals could be achieved within the proposed study ranges. Table 1 presents the results for 3 of the four critical responses for these runs. The results indicated that adjusting the factors within the proposed study ranges may enable meeting Goals 1 – 3: two runs yielded peak areas above 800, two runs resulted in a % RSD below 1.00, and all runs had migration times below 10 minutes.

Table 1. Tuning Run Results

Run No.	Peak Area – Mean	Peak Area – % RSD	Migration Time – Mean
1	897	5.08	8.53
2	643	2	5.55
3	431	1.4	3.99
4	148	5.73	4.22
5	464	2.52	6.26
6	830	0.55	7.7
7	526	0.13	4.04

In the second phase of the experimental work the Fusion QbD® software from S-Matrix Corporation (Eureka, CA USA) was used to expand the seven tuning runs into a 34-run model-robust optimization design. Fusion QbD model-robust designs have the advantage of visualizing nonlinear response trends and complex variable interaction effects that are not evident in data from classical factorial designs. Table 2 presents the model-robust experiment design, which consisted of the seven tuning runs (shown in blue background) and 27 added runs. As with the tuning runs, the Mean and % RSD response data are computed from three repeat runs for each experiment design run.

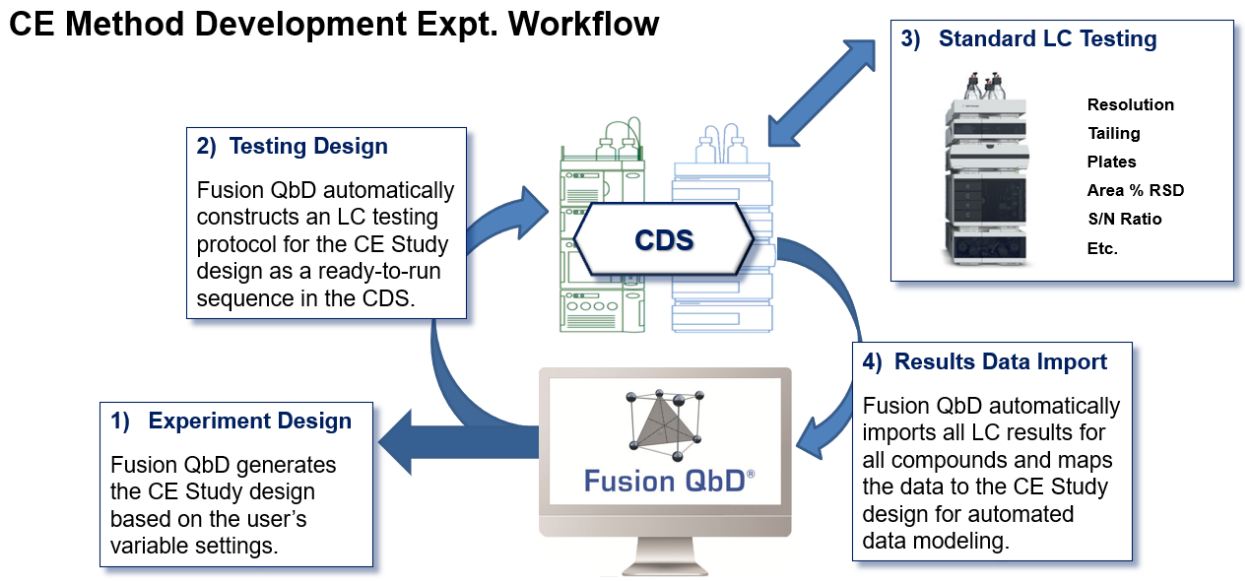
Table 2. – Model-robust Optimization Design Results

Run No.	Peak Area – Mean	Peak Area – % RSD	Migration Time – Mean
1	897	5.08	8.53
2	643	2	5.55
3	431	1.4	3.99
4	148	5.73	4.22
5	464	2.52	6.26
6	830	0.55	7.7
7	526	0.13	4.04
8	406	4.32	3.99
9	181	6.34	5.01
10	498	0.2	3.38
11	222	8.49	5.33
12	360	2.82	2.81
13	275	25.38	5.82
14	4,239	0	5.39
15	753	0.81	4.61
16	366	0.88	6.13
17	573	2.07	3.65
18	958	2.93	5.37
19	2,902	0	2.81
20	142	6.45	3.79
21	586	4.45	3.42
22	4,672	0	5.11
23	165	17.28	3.12
24	148	3.13	3.49
25	132	10.71	3.49
26	294	3.75	3.63
27	175	3.18	4.07
28	2,484	0	2.65
29	304	56.82	3.89
30	286	6.44	3.01
31	696	4.96	4.23
32	2,939	0	5.43
33	433	4.94	3.74
34	496	5.69	3.03

EXPERIMENT AUTOMATION

Figure 1 illustrates the experiment automation data flow between Fusion QbD and the CDS supporting CE method development experiments. Once the experiment design is generated in Fusion QbD (Step 1), a companion LC testing protocol is then generated in Fusion QbD and exported to the CDS as a ready ready-to-run sequence to which Fusion QbD attached the assay method (Step 2). The testing protocol contained all of the replicate injections required for each experiment design run, and the required standards protocol was also built into the sequence. Once the sequence is run and the chromatograms are processed in the CDS (Step 3), Fusion QbD automatically imports all results from the experiment chromatograms and maps them to the study for automated analysis and modeling (Step 4).

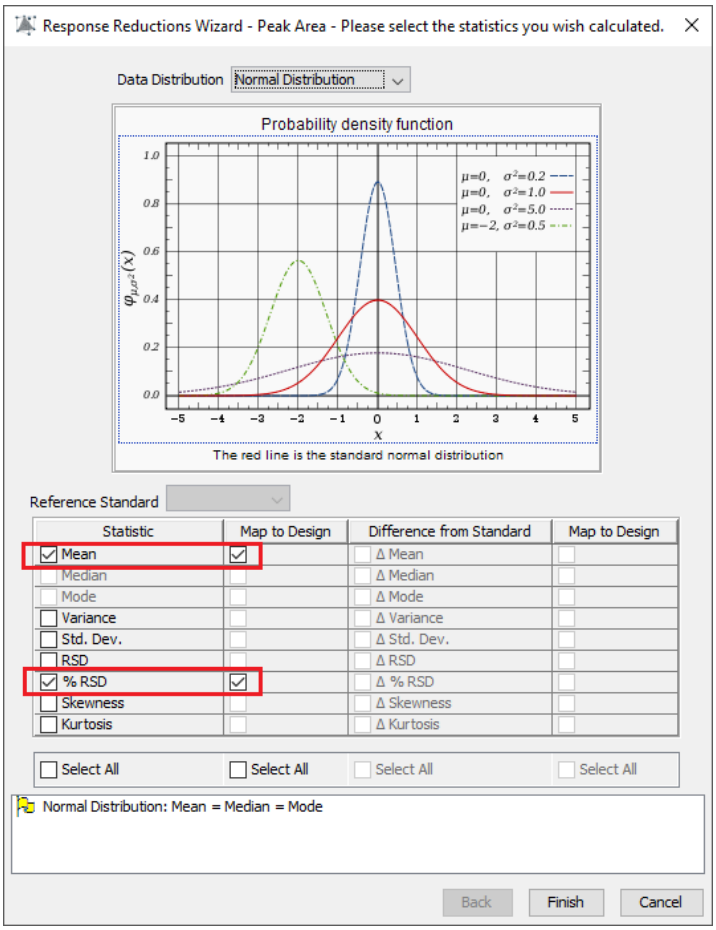
Figure 1. Fusion QbD Experiment Automation Support for the CE Study



COMPLEX DATA HANDLING AUTOMATION

Fusion QbD has a complex data handling capability to support multiple types of results data beyond the simple one-measurement-one result type of data. For example, Fusion QbD can automatically generate a variety of statistical data metrics from repeat test results per run, as shown in Figure 2. In this study Fusion QbD automatically generated the Mean ($\bar{X}$) and % RSD results from the replicate data for each run and mapped these statistical data reduction results to the experiment design runs for automated modeling.

Figure 2. Fusion QbD Complex Data Automation Support



Along with support for response data sets consisting of multiple test results per run, Fusion QbD also supports Time Series responses which consist of multiple time points per run. Fusion QbD can automatically create response profile curves and extract the critical results data needed for analysis and modeling. Additionally, Fusion QbD supports cascade impactor test results for R&D work on Orally Inhaled and Nasal Drug Products (OINDP), for which Fusion QbD can automatically generate all required particle size distribution results required for analysis.

RESULTS DATA MODELING

As Table 2 shows, the Peak Area – Mean response data range from a low of 132 to a high of 4,672. However, all but five of the 34 results (shown with yellow background) are below the lower acceptability limit of 1,000 area counts. These extremely high values were initially suspected to be outliers. However, as presented below, Fusion QbD accurately modeled this highly nonlinear response, and the customer was therefore able to demonstrate that methods existed within the combined study ranges which would generate these very high area responses.

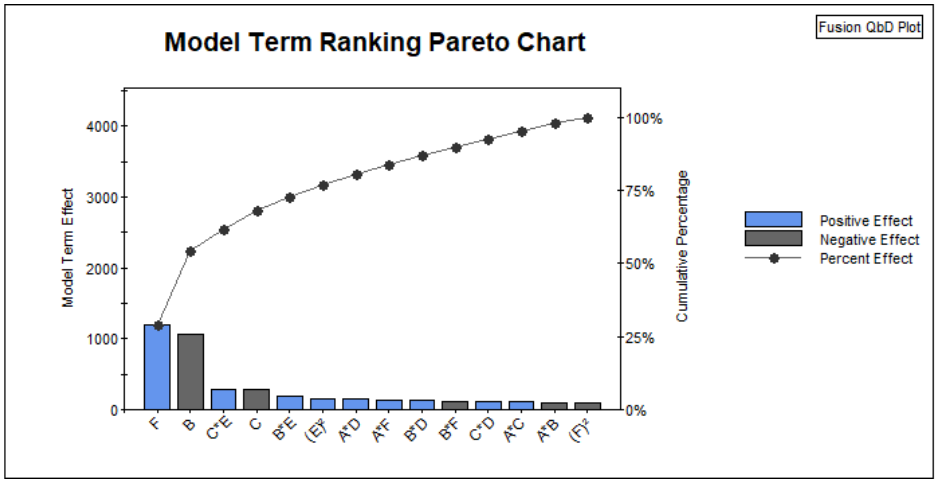
Fusion QbD automatically derived a model from the data which accommodated the nonlinear Peak Area response behavior. The model had an R^2 value of 0.9863, and Adjusted R^2 value of 0.9762, and a model standard error of ±230, which is of approximately the same magnitude as the overall experimental error of ±460 for this response as determined by the replicate run results. All terms in the Fusion QbD model are estimated to have statistically significant effects – i.e., effects with a magnitude statistically larger than can be attributed to experimental error.

Figure 3 presents a Pareto Ranking Chart of the study variable effects using the Fusion QbD model. As the chart shows, the effects of Vis. Agent (F) and BGE (B) are by far the dominant effects. However, their independent additive effects only account for 54% of the observed response variation (run-to-run differences in the response data set) – significant pairwise (two-variable) interaction effects and nonlinear effects involving all six variables are responsible for almost have of all observed data variation.

Figure 3. Peak Area Pareto Ranking Chart

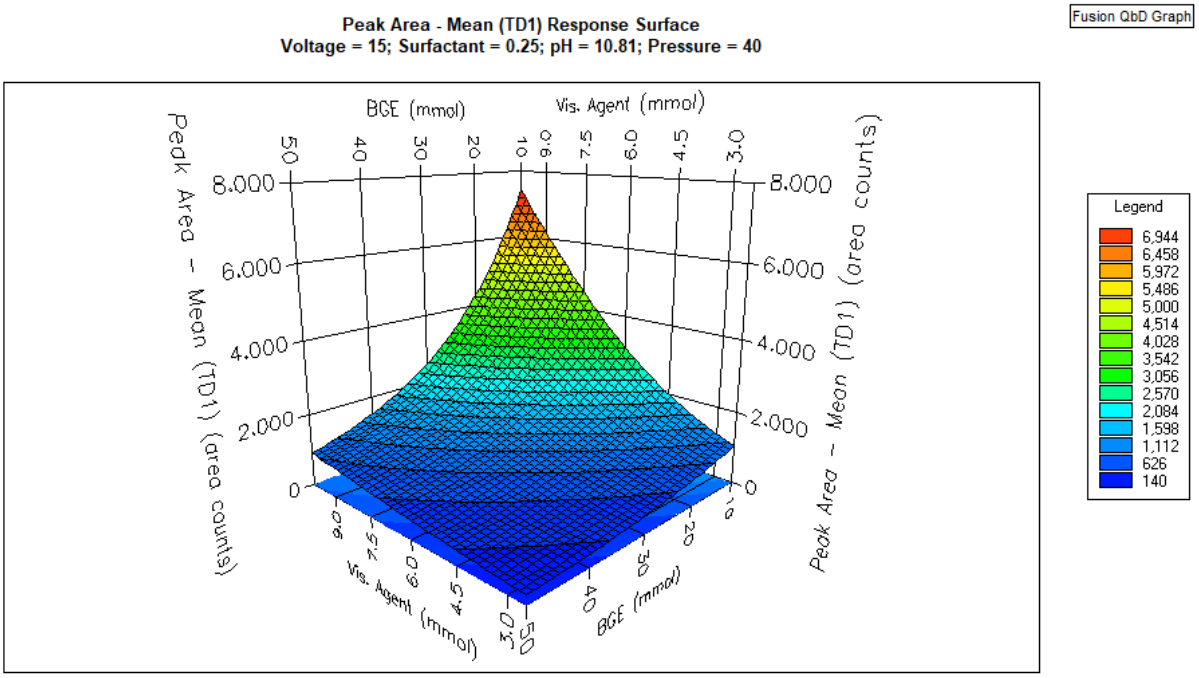
Study Variable Code Name Key

Study Variable	Units	Code Name
pH	*	A
BGE	mmol	B
Voltage	kV	C
Pressure	mBar	D
Surfactant	mmol	E
Vis. Agent	mmol	F



Interpreting the independent additive, interaction, and nonlinear effects of the study factors is best accomplished by graphical visualization using response surface (3D) graphs. Figure 4 is a response surface graph that simultaneously visualizes the combined effects of Vis. Agent and BGE on the Peak Area – Mean response. The graph shows that the ***main effect*** of BGE is slightly negative, i.e., *increasing* the BGE concentration *decreases* the mean Peak Area response, while the ***main effect*** of Vis. Agent is slightly positive - *increasing* the Vis. Agent concentration *increases* the mean Peak Area response. However, the dramatic pairwise interaction effect between these two parameters is also clearly evident in the graph – increasing the Vis. Agent concentration results in only a slight increase in Peak Area – Mean response at the high end of the BGE study range, while the increase results in a huge increase in the same response at the low end of the BGE study range.

Figure 4. Peak Area Response Surface



Fusion QbD provides numerical and graphical optimization capabilities to identify the study variable level settings that simultaneously meet user-defined goals for all responses included in the optimization search. A numerical optimization search was first carried out using its Best Overall Answer (BOA) search wizard which links the user-defined performance goals with the models derived from the data.

Table 3 presents the user defined performance goals entered into the wizard. Note that, as the Migration Time – Mean response was at or below five minutes for almost all runs, an upper acceptability limit of 5.00 minutes was used for this response in the optimization search. Note also that, as the API Resolution response values were substantially greater than 2.00 for all 34 runs, this response was not included in the optimization search.

Table 3. Best Overall Answer Search Goals

Response	Goal	Lower Acceptability Limit	Upper Acceptability Limit
Peak Area – Mean	Maximize	1000	
Peak Area – % RSD	Minimize		1.00
Migration Time – Mean	Minimize		5.0

Figure 5 presents the BOA search results based on the defined performance requirements. The figure presents the predicted overall best performing method obtained from the search, along with the corresponding predicted results for the included responses. The predicted mean responses were experimentally verified by running the predicted optimum CE assay settings several times at the predicted Best Overall Answer level settings of the six CE study factors

Figure 5. BOA Search Results

Variable Settings

Name	Level Setting	Units
pH	11.15	*
BGE	10	mmol
Voltage	21	kV
Pressure	30	mBar
Surfactant	0.25	mmol
Vis. Agent	10.0	mmol

Predicted Results

Response Name	Units	Goal	Predicted Result	Desirability	-2 Sigma Conf. Limit	+2 Sigma Conf. Limit
Peak Area - Mean (TD1)	area counts	Maximize	5,006	1.0000	2,972	8,473
Migration Time - Mean (TD1)	minutes	Minimize	3.66	0.2682	2.16	5.16
Peak Area - %RSD (TD1)	area counts	Minimize	0.04	0.9643	-0.10	1.79

Cumulative Desirability Target = 1.0000
Cumulative Desirability Result = 0.6371

The Acceptable Performance Region (APR) graphical optimizer was then used to examine how close the optimum setpoints of the study factors are to the edges of failure. Figure 6 is an Overlay graph generated by the APR graphical optimizer which shows how much of the combined study ranges of Vis. Agent and BGE, the two strongest effector variables, will meet or exceed the defined optimization goals when the remaining CE study factors are set to their BOA level settings. In these graphs each modeled response is assigned a color – in this case red for Peak Area – Mean, green for Peak Area – % RSD, and blue for Migration Time. The regions in Figure 6 shaded in a given color correspond to level setting combinations which **fail** to meet requirements for the response associated with that color. The remaining unshaded region in the lower right corner therefore represents level setting combinations of Vis. Agent and BGE which are predicted to meet or exceed the performance requirements defined for the graphed responses.

Figure 6. Graphical Optimizer – Multiple Response Overlay Graph

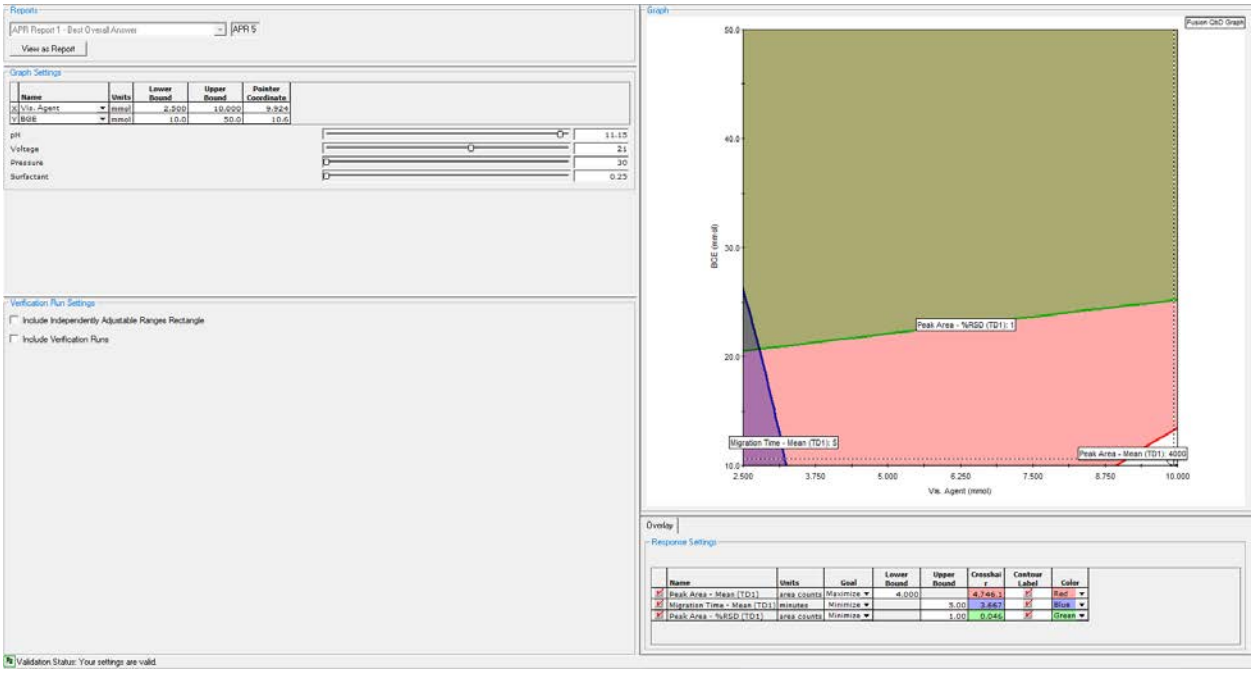


Figure 7 is a zoomed-in version of the previous APR graph obtained by reducing the graph ranges of Vis. Agent and BGE – the two graphed parameters. The graph in Figure 7 has also been adjusted by using the slider bars associated with the non-graphed parameters to identify the settings which result in the largest unshaded region. These refined settings become the new setpoints of the final method.

Figure 7. Adjusted APR Graph – Zoomed-in View

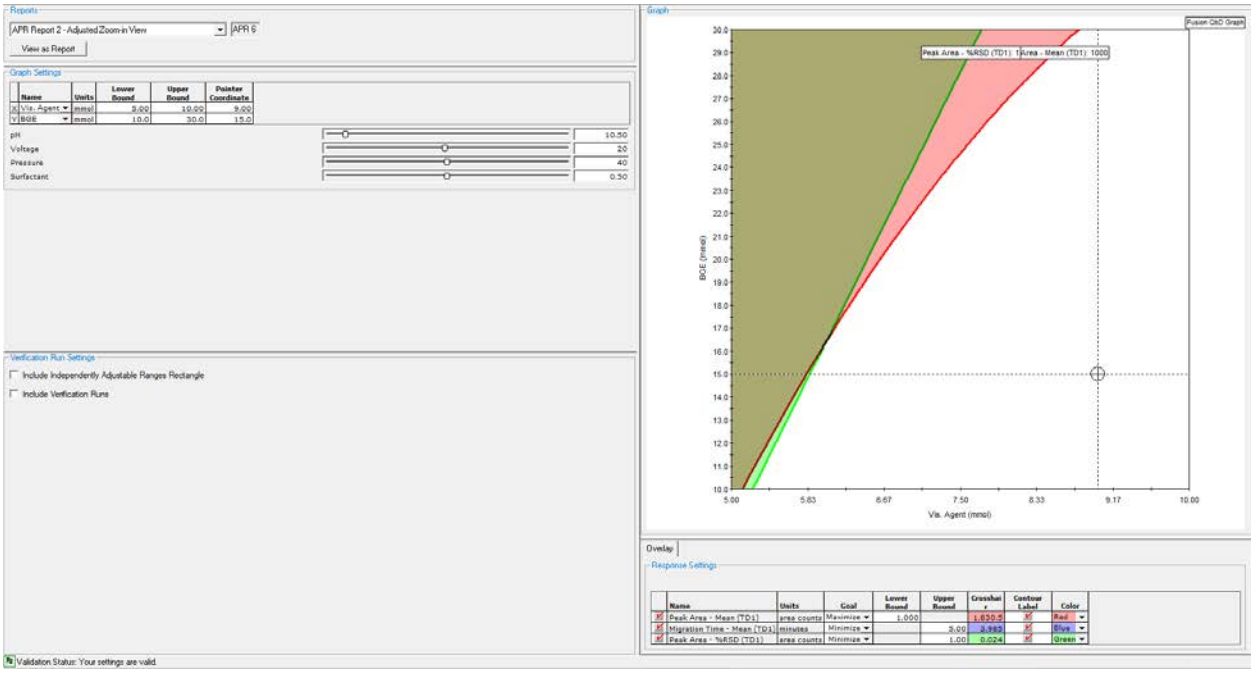
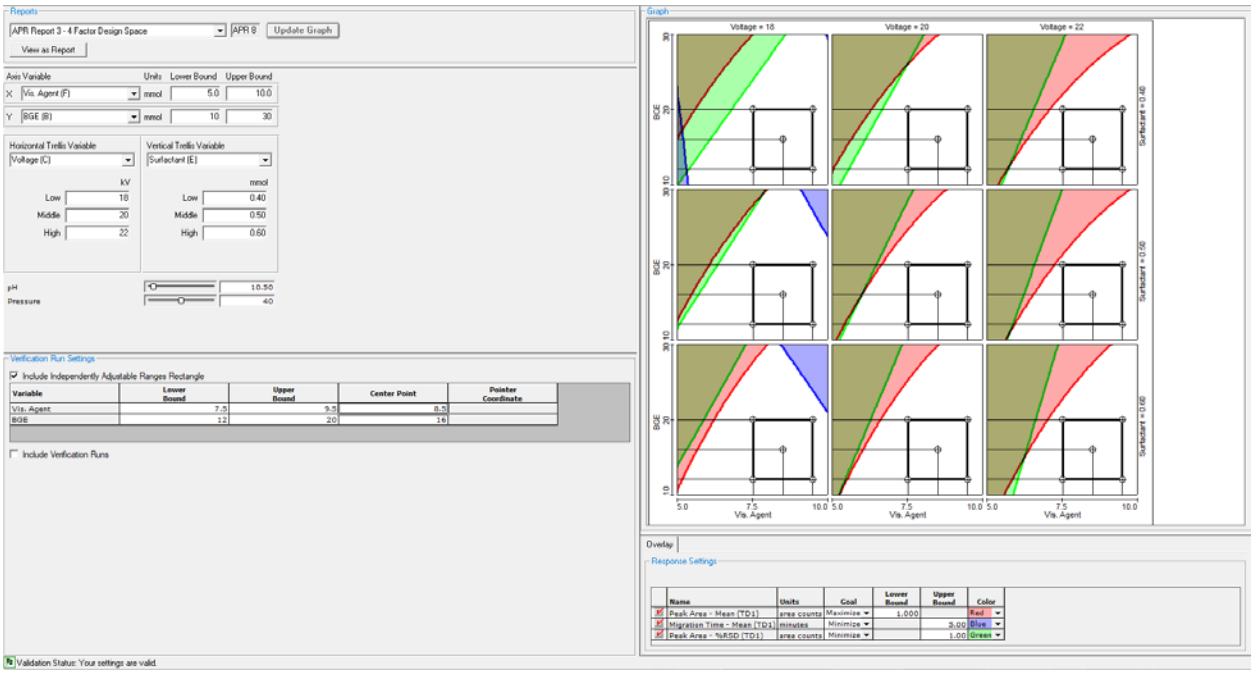


Figure 8 is a four-parameter design space graph series (a trellis graph series). The graph series includes graphs at the setpoints (Middle levels) and expected variation ranges (Low and High levels) of Voltage and Surfactant, the next two strongest effectors. Note that the middle graph in the nine-graph series has the same setpoints for Voltage, Surfactant, pH, and Pressure as the graph presented in Figure 7, and so the middle graph in the 3x3 graph trellis is the same graph previously presented in Figure 7.

Figure 8. Four Parameter Design Space



Fusion QbD enables an Independently Adjustable Range (IAR) rectangle to be superimposed on these graphs to designate allowable post-approval permanent changes to the graphed variables within the design space – these are changes which can be independently made while maintaining performance requirements for the modeled responses. In this case the rectangle is set to a range of 7.50 – 9.50 for Vis. Agent (final method setting = 8.50) and 12.0 – 20.0 for BGE (final method setting = 16). Figure 8 also shows that the two trellis graph variables (Voltage = horizontal trellis series, Surfactant = vertical trellis series) have safe operating ranges of 18.0 – 22.0 for Voltage (final method setting = 20.0) and 0.40 – 0.60 for Surfactant (final method setting = 0.50). Note that all graphs in the trellis correspond to final method settings of 10.50 for pH and 40.0 for Pressure.

CONCLUSIONS

The combination of a QbD aligned statistical experimental design approach with rigorous, expert-system modeling within Fusion QbD enabled accurate quantification of all the important CE study parameter effects on the critical responses. In this study these effects included both highly nonlinear and highly interactive study parameter effects. This quantification identified a robust design space and a final method which exceeded all performance requirements, thereby achieving the goal of exchanging a costly HPLC assay for a quick and economical CE assay.

Note – disparate results such as shown by the five high response value in Table 2 are often prematurely dismissed as outliers, even when the disparity is in an advantageous direction. Outliers are normally an expression of one of three circumstances: an assignable cause of incorrect run execution, such as a mistake in weighing or setting a study factor to the wrong level for a run; a change in an unidentified effector such as batch, calibration, or humidity; or highly nonlinear and/or highly interactive study parameter effects, as in this case. This is why apparent outliers should always be investigated for an assignable cause, and also why predicted optimum results should always be experimentally verified.