Supplementary Material

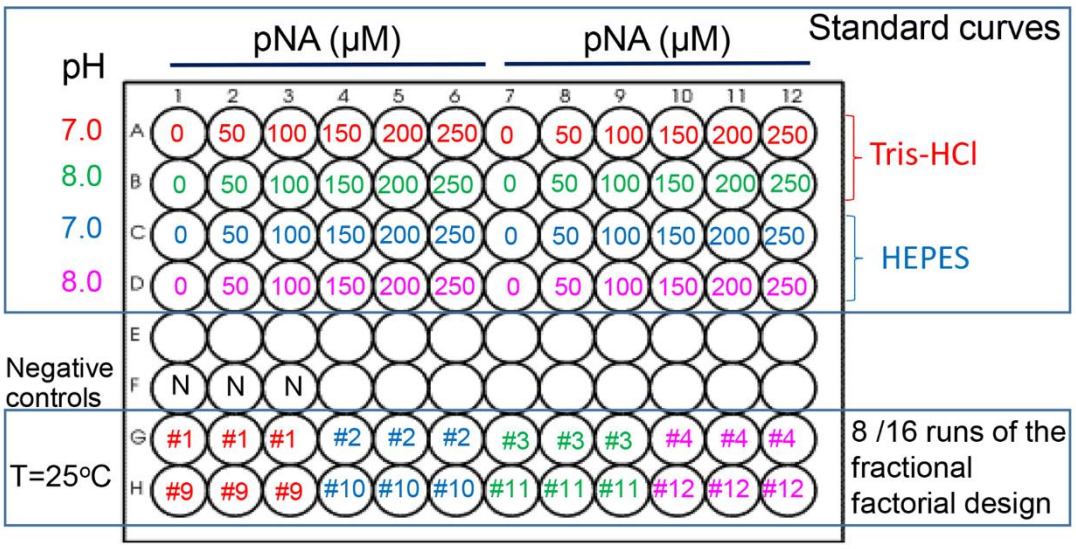
A general guide for the optimization of enzyme assay conditions using the design of experiments approach

by

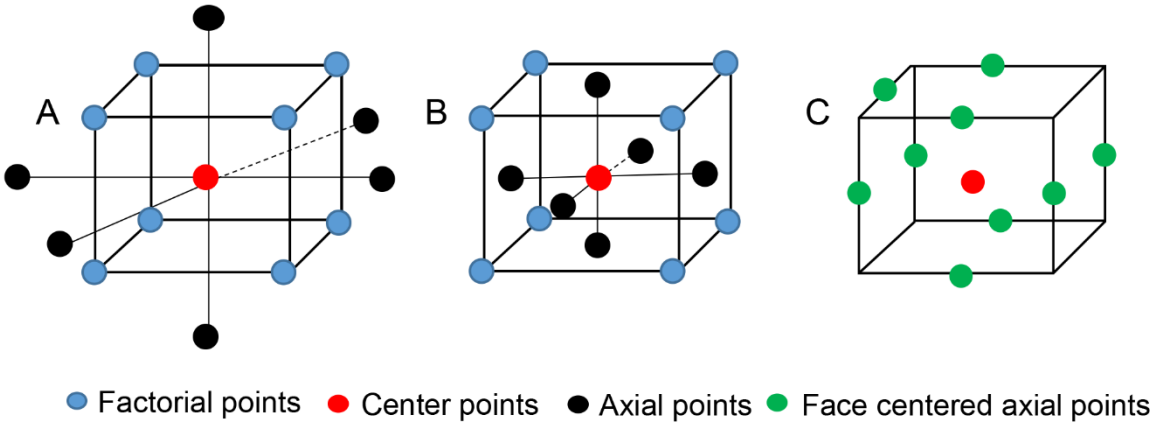
Favour Chinyere Onyeogaziri and Christos Papaneophytou*

Department of Life and Health Sciences, School of Sciences and Engineering, University of Nicosia, 46 Makedonitissas Ave., 1700, Nicosia, Cyprus

*Correspondence: Christos Papaneophytou, Department of Life and Health Sciences, School of Sciences and Engineering, University of Nicosia, Nicosia, Cyprus email: papaneophytou.c@unic.ac.cy; Tel: +357 22 841941



Supplementary Figure S1: Example of pipetting scheme for assessment of HRV-3C protease activity under various reaction conditions with a triplicate reaction for each sample. A different *p*NA standard curve was developed for each of the various buffers that were tested. Note that the different buffers are color-coded.



Supplementary Figure S2: Graphic representations of central composite design (CCD) (**A**), face-centered CCD (**B**) and Box-Behnken design (**C**). When an axial distance of 1.0 is chosen then the CCD has only 3 levels per factor and is called the Face-Centered Design.

Supplementary Table S1 Vocabulary of DoE

Term	Description
Experimental Design	A set of experiments composed of a matrix of different level combinations of the factors studied
Factor	One of the tested (independent) variable that can be set to desire value
Categorical factor	A factor with non-numerical (e.g., buffer composition could be either PBS or TBS)
Continuous factors	A factor with numerical values (e.g., DTT concentration could be either 5 or 10 mM)
Level	The value (numerical or qualitative "label") of a factor
Run	An experiment under specific reaction conditions
Response	The numerical result of an experiment
Full factorial	All possible combinations of variables
Fractional factorial	Only a fraction of the full factorial design is carried out.

Supplementary Table S2 Essential terms that must be consider in analysis of variance

Term	Description
Model F Value	It compares the variance of a term with residual (error) variance. If the variances are close to the same, the ratio will be close to one and the term does not have a significant effect on the response (the higher the better).
Model *p*-value	It should be significant and < 0.05 to be strongly significant. When the model *p*-value is between 0.05 and 0.10 is slightly significant
Term *p*-values	Each of the terms should have a *p*-value < 0.05 to be significant. Sometimes insignificant terms (factors) are excluded from the model.
R-squared values	The most important are the "adjusted" and "predicted" R-squared values. Usually, an R-squared value above 0.6 is required
Adequate Precision	It is the signal to noise ratio and it should be >4. It compares the range of the predicted values at the design points to the average prediction error
Lack of fit	It should be insignificant (p>0.05). A significant lack of fit means that the design points are not fitted well to the polynomial model

Supplementary Table S3 ANOVA results for response surface quadratic model for HRV-3C protease activity

Source	Sum of Squares	df	Mean Square	F Value	P value*	
Model	3.122005×10^{5}	14	22297.90	16.15	< 0.0001	Significant
A (pH)	38203.10	1	38203.10	27.67	< 0.0001	Significant
B (Time, min)	17421.12	1	17421.12	12.62	0.0029	Significant
C (DTT, mM)	55978.99	1	55978.99	40.54	< 0.0001	Significant
D (Glycerol, %v/v)	2283.66	1	2283.66	1.65	0.2179	Not significant
AB	154.68	1	154.68	0.11	0.7425	Not significant
AC	7.69	1	7.69	5.566×10^{-3}	0.9415	Not significant
AD	0.72	1	0.72	5.18710^{-4}	0.9821	Not significant
BC	2239.64	1	2239.64	1.62	0.2222	Not significant
BD	8156.97	1	8156.97	5.91	0.0281	Significant
CD	465.97	1	465.97	0.34	0.5699	Not Significant
A^2	27083.87	1	27083.87	19.61	0.0005	Significant
B^2	2271.62	1	2271.62	1.65	0.2191	Not significant
C^2	1.432005×10^{5}	1	1.432×10^{5}	103.70	< 0.0001	Significant
D^2	34259.56	1	34259.56	24.81	0.0002	Significant
Residual	20711.99	15	1380.80			
Lack of Fit	17209.33	10	1720.93	2.46	0.1666	Not significant

*p value: $p<0.05$ is consitered as significant

R^2= 0.9378; Adj R^2 =0.8797 ; Pred R^2=0.6871; Adeq Precision= 18.111