Journal of Chromatography A, 1223 (2012) 64–71

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

High productivity chromatography refolding process for Hepatitis B Virus X (HBx) protein guided by statistical design of experiment studies

Anindya Basu, Susanna Su Jan Leong*

School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, Singapore

ARTICLE INFO

Article history:
Received 2 September 2011
Received in revised form 8 December 2011
Accepted 8 December 2011
Available online 16 December 2011

Keywords:
Hepatitis B Virus X protein
Chromatography refolding
Immobilised metal affinity chromatography
Design of experiments
Bioprocess intensification

ABSTRACT

The Hepatitis B Virus X (HBx) protein is a potential therapeutic target for the treatment of hepatocellular carcinoma. However, consistent expression of the protein as insoluble inclusion bodies in bacteria host systems has largely hindered HBx manufacturing via economical biosynthesis routes, thereby impeding the development of anti-HBx therapeutic strategies. To eliminate this roadblock, this work reports the development of the first 'chromatography refolding'-based bioprocess for HBx using immobilised metal affinity chromatography (IMAC). This process enabled production of HBx at quantities and purity that facilitate their direct use in structural and molecular characterization studies. In line with the principles of quality by design (QbD), we used a statistical design of experiments (DoE) methodology to design the optimum process which delivered bioactive HBx at a productivity of 0.21 mg/ml/h at a refolding yield of 54% (at 10 mg/ml refolding concentration), which was 4.4-fold higher than that achieved in dilution refolding. The systematic DoE methodology adopted for this study enabled us to obtain important insights into the effect of different bioprocess parameters like the effect of buffer exchange gradients on HBx productivity and quality. Such a bioprocess design approach can play a pivotal role in developing intensified processes for other novel proteins, and hence helping to resolve validation and speed-to-market challenges faced by the biopharmaceutical industry today.

1. Introduction

The Hepatitis B Virus X protein (HBx) is envisaged to be a potential therapeutic target for the treatment of hepatocellular carcinoma (HCC) which affects millions of lives every year [1,2]. The current medications approved for HCC within the USA have been rather ineffective in ensuring a complete cure and pose the threat of development of viral resistance and serious side effects amongst the treated patients [3,4]. There is therefore an urgent need for the development of new therapeutic strategies. HBx, being a multi-functional viral regulator, has been well-established as a causative factor for host cellular transformations, which makes it an optimal therapeutic target [5–7]. However, the inherently low HBx expression in the infected host cells and the incessant insoluble expression of the protein in microbial expression systems pose a tremendous challenge for HCC scientists involved in structural characterization and subsequent drug designing studies of HBx. This challenge forms the basis of this study which aims to establish good understanding of how different bioprocess parameters and their interaction affect HBx refolding, leading to the rational development of an intensified bioprocess for HBx production.

We chose to develop a HBx bioprocess based on immobilised metal affinity chromatography (IMAC), which enables simultaneous purification and refolding of the 6His-tagged protein, thus minimising the number of unit operations involved to improve productivity and process cost [8,9]. Bioprocess studies for proteins like HBx, which has no commercially available native standards, demands a methodical and rational approach in process development [10]. Instead of adopting a trial and error approach for optimising the IMAC refolding process, we employed a statistical design of experiment (DoE) methodology to extract maximum information about the impact of different refolding process parameters on the quality of the HBx product. Our systematic DoE based approach for designing and optimising the IMAC refolding process are aligned with the requirements of drug regulatory authorities to implement quality by design (QbD) principles in biopharmaceutical manufacturing operations [11,12]. Although several IMAC refolding processes have been reported in the literature for various model proteins [17,18], the interaction effects of the different process parameters on the critical to quality attributes (CQA) of the product has not been systematically studied, which renders the need for each IMAC process to be optimised empirically for different proteins. Using this DoE-directed

* Corresponding author. Fax: +65 6794 7553.
E-mail address: SLeong@ntu.edu.sg (S.S.J. Leong).

0021-9673/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2011.12.037

approach to design and optimise the IMAC refolding process, an intensified IMAC-based bioprocessing platform which enabled the production of bioactive HBx at ~350 mg/L and >95% purity in a streamlined process flowsheet was successfully developed. We expect this IMAC bioprocess platform to open the way for scalable production of bioactive HBx proteins, which can also be readily extended for refolding of other Histidine-tagged model proteins.

2. Methods and materials

All chemicals were obtained from Sigma–Aldrich (Singapore), unless stated otherwise.

2.1. HBx Protein expression and solubilisation

The gene sequence for the HBx protein, from the HBV subtype A3 (Isolate Cameroon), was cloned into a pET28b+ vector and was used for expression in *E. coli* BL21(DE3)RIL cells (Stratagene, Singapore). Expression and purification of the inclusion bodies (IB) containing the HBx protein were carried out using the same methods described in our earlier work [13]. The prepared HBx IBs were solubilised in a denaturing–reducing buffer (8 M urea, 1 M NaCl, 10 mM DTT, 20 mM imidazole, 50 mM Tris, pH 7.4) at 0.8–1.2 mg/ml protein concentration for 3 h. The dissolution buffer composition used to bind denatured-reduced HBx on an IMAC column was designed to bind the protein at high amounts and selectivity. 8 M urea helps to denature/unfold the protein, 1 M NaCl minimises non-specific electrostatic interaction on the column resins, 10 mM DTT is added to reduce HBx disulfide bonds which will interfere with further refolding, a low concentration of Imidazole (20 mM) is added to prevent non-specific binding of contaminant proteins containing aromatic residues. The dissolution time was optimised during our preliminary studies; the minimum time required for homogenous dissolution of the IBs was found to be 3 h, as determined by RP-HPLC analysis. The denaturing-reducing buffer used for solubilising the HBx protein was also used for pre-equilibrating the IMAC columns and would be referred to as IMAC binding buffer in the following sections.

2.2. IMAC refolding of the HBx protein

Refolding studies for the HBx protein were performed on IMAC columns using the AKTA Explorer Fast Protein Liquid Chromatography (FPLC) system (GE Healthcare, Singapore). The process flowsheet for IMAC refolding and purification of HBx is shown in Scheme 1 (Fig. 1). All experiments were performed using 1 ml Ni^{2+} Sepharose HisTrap HP columns (GE Healthcare, Sweden) at 1 ml/min flow-rate. 5–10 mg of solubilised denatured–reduced IBs (as determined by Bradford assay and gel densitometry analysis) were loaded on the Ni^{2+} Sepharose columns which were pre-equilibrated with IMAC binding buffer. After protein loading, the column mobile phase was changed from IMAC binding buffer to refolding buffer (2 M urea, 0.25 M Arginine, 0.1 mM GSH and 0.01 mM GSSG in 50 mM Tris, pH 7.4) over 0–5 column volumes (CVs), followed by equilibration with 15 CV of refolding buffer. The Ni^{2+} Sepharose columns containing the bound proteins were then incubated at 4 °C for different time periods (1–3 days), and the proteins were then eluted in a single step by the refolding elution buffer (2 M urea, 0.25 M Arginine, 0.1 mM GSH and 0.01 mM GSSG, 1 M Imidazole in 50 mM Tris, pH 7.4), followed by determination of the refolding yield. The column was then washed with 10 CV of IMAC binding buffer and a stripping step was performed using the stripping buffer (8 M urea, 1 M NaCl, 10 mM DTT, 1 M imidazole, 50 mM Tris, pH 7.4) to remove strongly bound proteins.

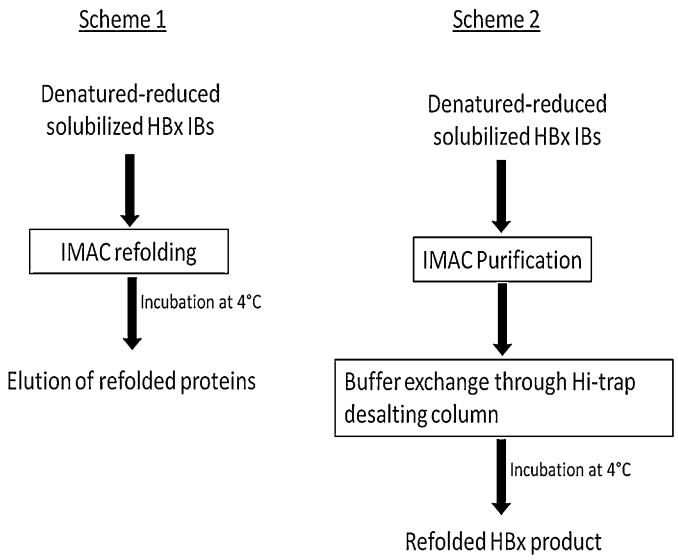

Fig. 1. HBx refolding process flowsheets based on (1) IMAC refolding, and (2) dilution refolding.

2.3. Dilution refolding of HBx IBs

The process flowsheet for HBx dilution refolding and purification is shown in Scheme 2 (Fig. 1). The prepared IBs were first purified with a 1 ml Ni^{2+} Sepharose HisTrap HP column using the AKTA Explorer FPLC system at a constant mobile phase flow rate of 1 ml/min. 10 mg of the solubilised denatured–reduced HBx IBs were loaded into a Ni^{2+} Sepharose column which was pre-equilibrated in the IMAC binding buffer. The bound proteins were then eluted with the stripping buffer. Dilution refolding was performed in a Hi-Trap Desalting column (GE Healthcare, Sweden) which was equilibrated in the refolding buffer (2 M urea, 0.25 M Arginine, 0.1 mM GSH and 0.01 mM GSSG in 50 mM Tris, pH 7.4). The purified HBx IBs were loaded into the desalting column, and the eluate which was 1.33-fold diluted was collected, incubated overnight at 4 °C, and the HBx refolding yield determined.

2.4. Analytical methods

Protein concentrations were measured by Bradford assay.

2.4.1. Determination of HBx soluble yields

HBx soluble yield and purity were measured by a chip-based electrophoresis method performed on an Agilent 2100 Bioanalyzer® (Agilent, Singapore) in combination with the Protein 230 LabChip® kit. Protein signal detection within this instrument is based on laser-induced fluorescence of an intercalating dye interacting with protein–SDS complex which is then analysed by the Agilent 2100 Bioanalyzer software. Due to the absence of any commercially available native HBx standards, for each chip based analysis, the solubilised HBx IBs (concentrations determined by Bradford assay) were diluted to different concentrations and used to generate a calibration curve for determining HBx concentrations after dilution and IMAC refolding. The samples and chips were prepared according to the protocol provided with the Protein 230Plus LabChip® kit. All samples used for the analysis were centrifuged and filtered using a 0.2 μm filter prior to analysis. HBx soluble yields were determined by Eq. (1):

Table 1
Independent (process) variables and the corresponding coded values used for the DoE study.

Factors	Coded symbol	Actual values of the coded levels		
		−1	0	1
Incubation period (days)	X_1	1	2	3
Protein load (mg)	X_2	5	7.5	10
Buffer exchange gradient (CV)	X_3	0	2.5	5

$$\text{Soluble yield} = \frac{\text{Mass of soluble HBx as determined by chip-based electrophoresis}}{\text{Mass of denatured - reduced HBx loaded}} \times 100\% \quad (1)$$

2.4.2. Determination of HBx refolding yields

HBx protein refolding yields were determined by an ELISA kit developed in-house for exclusively determining the refolding yields of 6His-tagged HBx proteins [14]. This ELISA system is based on the ability of bioactive HBx protein molecules to interact with the GST-tagged p53 protein ligand. The GST-p53 molecules were immobilised on a GSH functionalised maleimide surface to avail the p53 ligand to bind with bioactive HBx molecules present. The amount of refolded HBx molecules bound to the p53 surface was determined by measuring the absorbance signal generated by the interaction between bound 6His-HBx and Horse Radish Peroxidase (HRP)-tagged 6His primary antibody (Abcam, USA). Since a maleimide surface can specifically bind to proteins containing free sulfhydryl groups, a calibration curve needed to determine the amount of p53-bound bioactive HBx was generated by incubating standard 6His-GST proteins (immobilised directly on the maleimide functionalised surface under mild reducing conditions) with the HRP-tagged 6His primary antibody. As reported in our previous study, the validity of the ELISA assay was demonstrated by the low coefficient of variation (<6%) for all the tested samples. HBx refolding yields were calculated based on Eq. (2).

$$\text{Refolding yield} = \frac{\text{Mass of bioactive HBx as determined by ELISA}}{\text{Mass of denatured - reduced HBx IBs loaded}} \times 100\% \quad (2)$$

2.5. Development of an optimised IMAC refolding process using statistical DoE

The design of the IMAC refolding process was guided by the statistical design of experiments (DoE) methodology. The impact of different process parameters on our product CQA (i.e. HBx protein yields and concentration) were determined by an Ishikawa diagram, which helps to identify all the parameters affecting the CQA of the process [15,16]. Optimisation of the IMAC refolding process was performed by employing a two level full factorial central composite experimental plan with three operating parameters as the independent variables: (i) on-column incubation period (number of days), (ii) protein load, and (iii) the rate of change of denaturing buffer to refolding buffer (number of CVs used to perform the buffer exchange). The levels of the factors considered and the corresponding coded symbols used for the DoE analyses are summarised in Table 1. The experimental plan consisted of 25 trials including one centre point as detailed in Table 2. The response variables (i.e. HBx protein yield and concentration) were fitted into a polynomial equation (Eq. (3)):

$$Y_i = \beta_o + \sum \beta_i x_i + \sum \beta_{ij} x_i x_j + \sum \beta_{ijk} x_i x_j x_k \quad (3)$$

Table 2
DoE experimental design for HBx IMAC refolding and corresponding results.

Coded levels[a]			Soluble yield (%)	Concentration (mg/ml)
X_1	X_2	X_3		
−1	+1	−1	55	0.5
+1	−1	+1	20	0.1
−1	−1	+1	70	0.3
+1	+1	−1	40	0.37
−1	−1	−1	20	0.07
+1	+1	+1	21	0.2
+1	−1	−1	20	0.15
+1	−1	−1	18	0.12
+1	+1	+1	25	0.2
−1	+1	−1	60	0.55
−1	+1	+1	65	0.65
+1	+1	+1	20	0.2
−1	+1	−1	60	0.54
−1	−1	−1	25	0.08
−1	+1	+1	61	0.61
+1	+1	−1	40	0.4
+1	−1	+1	24	0.12
−1	−1	+1	75	0.32
−1	−1	−1	20	0.09
−1	−1	+1	75	0.35
0	0	0	43	0.315
+1	−1	+1	25	0.13
−1	+1	+1	62	0.62
+1	−1	−1	15	0.13
+1	+1	−1	45	0.35

[a] Details of symbol explained in Table 1.

where Y_i is the predicted response variable, while $x_i x_j$ are the independent variables. β_o is the offset term, β_i is the ith linear coefficient for the main effects, β_{ij} is the coefficient for the two-way interactions while β_{ijk} is the interaction coefficient for the three-way interactions. The polynomial coefficients were calculated and analysed using the 'Minitab 15' (Minitab Inc., State College, PA, USA) statistical software package. Statistical analyses of the models were performed by the analysis of variance (ANOVA).

3. Results and discussion

3.1. DoE scheme and data analysis method for HBx bioprocess development

The DoE approach adopted to guide the development of the HBx IMAC refolding process is summarised in Fig. 2. The overall aim for employing the DoE approach is to systematically (i) determine the effects of different variables or process parameters of the IMAC refolding process on the process CQA (i.e. HBx refolding yield and concentration), and (ii) obtain the design space which can provide the optimum CQAs for the target product, HBx. An Ishikawa diagram was constructed to identify different process parameters that can directly impact the process CQA for the DoE analysis [12,15]. Based on the number of parameters selected, a set of experiments were then designed at statistically determined points to obtain maximum information with minimum number of experiments. The experimental design, in general, largely depends on the number of parameters to be tested and the feasibility of the experimental set-up required or the resolution desired [17]. In this study, a two-level full factorial central composite design was used along with one centre point to obtain full resolution data for the HBx refolding process. A standardised effects plot was then constructed to identify which of the parameters, or their combined interactions, significantly affected the response variables. The terms that significantly influenced the response variables (or the CQAs) were then chosen to develop semi-empirical models which would provide the predicted responses of the CQA parameters within the tested range of

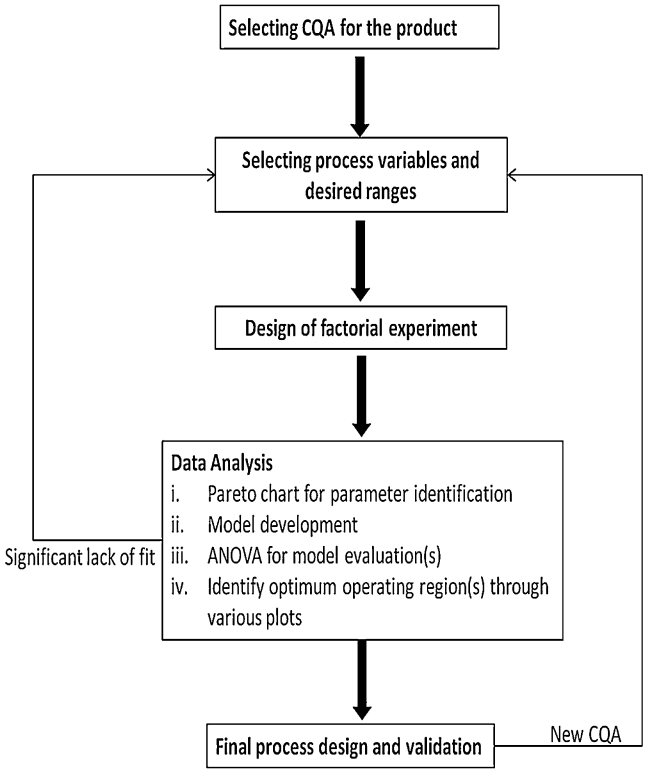

Fig. 2. DoE scheme employed for rational development of a HBx refolding process.

the process variables. Statistical quality of the models was assessed by ANOVA studies which can provide information on: (a) the significance of the parameters used within the models, and (b) the source of residual errors obtained from the models. In general, the residual error possesses two important components: lack of fit (error due to improper model fitting) and pure error (experimental error). Any critical process parameters if excluded from the study would result in a significant lack of fit error ($P < 0.05$), necessitating a reconsideration of the process variables tested. Graphical plots from the developed models were then used to determine the optimum process design which maintains the CQA parameters within the desired range. The model-predicted outcomes were subsequently validated with experimental results.

3.2. Selection of IMAC refolding parameters for DoE studies

The aim of this study is to develop a bioprocess that would allow the production of HBx protein at concentrations and purity which enable their direct use in structural and drug designing studies. To this end, the two most important CQA requirements that the process must achieve are (i) high HBx refolding yield, and (ii) high HBx protein concentration. To accelerate data acquisition in the optimisation study of IMAC refolding parameters, HBx soluble yields were measured instead of HBx refolding yield for preliminary screening of refolding conditions to indicate correct refolding. This decision was made based on the results of our earlier study, where HBx soluble yield was found to positively correlate with HBx refolding yield [14]. Based on the DoE approach employed, as discussed in Section 3.1, three process parameters from a large number of parameters that were considered to be able to significantly affect the IMAC refolding process (Fig. 3) were chosen for optimisation studies.

Since all column refolding experiments were performed on prepacked 1 ml Ni^{+2} Sepharose columns (with fixed dimensions) at a fixed flow rate, process parameters such as column equilibration, bed height and flow rate were not chosen for optimisation in this study. Protein load was chosen as the first process variable to be studied, where protein binding behaviour on-column has been reported to significantly affect refolding yield [18–20]. We hypothesised that a higher protein load may reduce protein intermolecular distances leading to unwanted protein aggregation during on-column refolding. The total mass of protein loaded was thus chosen as a process parameter for our DoE study. The highest protein concentration used in previously reported IMAC refolding studies was approximately 5 mg protein per ml of resin [20]. Since the aim of this study is to develop an intensified refolding platform to increase HBx productivity, a high protein load is desired. The effects of protein concentration in the range of 5–10 mg/ml on HBx refolding yield and concentration was therefore chosen to be studied.

The success of any refolding process relies on the use of an optimal refolding buffer that gives maximum refolding yield. We previously reported a second virial coefficient (SVC) based methodology for obtaining the optimal refolding buffer for dilution refolding of the HBx protein [13,21]. The same refolding buffer was used in this study. Upon immobilisation of the denatured–reduced HBx protein molecules on the Ni^{2+} resin, successful protein refolding would depend on the flexibility of the HBx polypeptide to refold within the physicochemical environment of the mobile phase [9,22]. We hypothesise that refolding will be influenced by the rate of change of mobile phase from denaturing to refolding, which would largely depend on rate of mobile phase buffer exchange from denaturing to refolding [17,23,24]. Buffer exchange gradients were hence chosen as another process variable for the DoE study. With the aim to minimise refolding time for increased productivity, buffer exchange gradients from 0 to 5 CV were employed for our IMAC refolding studies. In our earlier dilution refolding studies, we observed that maximum refolding yields of the HBx protein (~55% at a refolding concentration of 0.1 mg/ml) were obtained only after 3 days of protein incubation at 4 °C [14], which indicates a relatively slow refolding kinetics for the HBx protein. Hence, in this work, the incubation period (number of days) was chosen as the third process variable for the DoE study. The maximum incubation time under refolding conditions was kept at 3 days. To maximise protein concentration during elution, step elutions rather than gradient elutions were employed for all the IMAC refolding experiments. The process variables (independent variables) chosen for the DoE along with the levels and coded symbols are presented in Table 1.

3.3. Development and optimisation of an IMAC refolding bioprocess for HBx production

As per the DoE strategy, three process variables were studied to optimise HBx refolding on the IMAC column. The impact of different combinations of the process variables on HBx soluble yield and the final protein concentration is summarised in Table 2. Preliminary ANOVA studies of the experimental data did not show any significant contribution to data variability due to curvature or non-linearity ($P > 0.05$) of the response variables within the limits of the tested independent variables (data not shown). Since the centre point (X_1, X_2, $X_3 = 0$, i.e. incubation time = 2 days, protein load = 7.5 mg and buffer exchange gradient = 2.5 CV) within our DoE design is representative of the curvatures within the experimental range, we excluded this point during our subsequent data analyses. Fig. 4 is a pareto chart which compares the importance of the different single parameters and also interaction between the parameters studied on HBx soluble yield and protein concentration during IMAC based on the results obtained in Table 2. It is clear that the protein incubation period has a major impact on both the soluble yield and concentration of the HBx protein. The parameters having a significance level higher than 95% (as indicated by the vertical line in Fig. 4) were then chosen for model development to

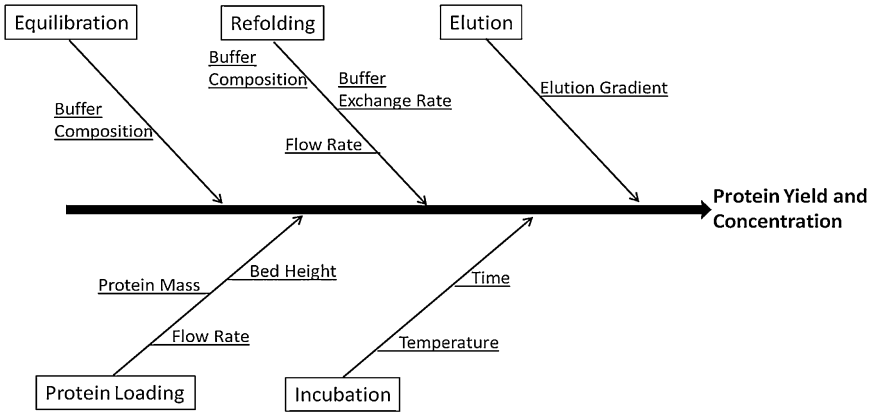

Fig. 3. Ishikawa diagram showing the different factors that can affect the performance of a column refolding process.

predict HBx soluble yield and protein concentration under varying conditions. The models developed for prediction of soluble yield and protein concentration after IMAC refolding are shown in Eqs. (4) and (5) below.

$$\text{Soluble yield}\ (\%) = 40.04 - 13.96X_1 + 6.13X_2 + 5.21X_3 - 8.79X_1X_3 - 9.04X_2X_3 + 2.79X_1X_2X_3 \quad (4)$$

$$[R^2 = 98.85\%;\ R^2(\text{pred}) = 97.70\%;\ R^2(\text{adj}) = 98.44\%]$$

$$\text{Concentration}\ (\text{mg/ml}) = 0.3 - 0.092X_1 + 0.135X_2 + 0.019X_3 - 0.054X_1X_2 - 0.066X_1X_3 - 0.038X_2X_3 \quad (5)$$

$$[R^2 = 99.31\%;\ R^2(\text{pred}) = 98.62\%;\ R^2(\text{adj}) = 99.06\%]$$

The quality of the developed models are indicated by the R^2, R^2(pred) and R^2(adj) values, where R^2 indicates the net variability within the data that is explained by the respective model. Therefore, the closer the value is to 100%, the better the model. The predicted R^2[R^2(pred)] indicates how well the model will describe any future data, while the adjusted R^2[R^2(adj)] is a modified form of R^2 that takes into consideration the number of terms used within the model. The models developed through this DoE strategy describe the experimental results obtained well (in Table 2), and could therefore be used to predict the optimum process conditions for the IMAC refolding strategy. Furthermore, ANOVA results for our developed models indicate that the error due to lack of fit is insignificant ($P > 0.05$) (Table 3), indicating that the variable process parameters chosen for our investigation was sufficient to explain the observed data variability seen in Table 2.

3.4. Designing the optimum IMAC refolding strategy and comparison with dilution refolding

With reliable predictive models (i.e. Eqs. (4) and (5)) developed for the HBx IMAC refolding process, our next aim was to determine the optimum refolding parameters for HBx refolding on-column. From Fig. 4A, it is clear that the most important parameter affecting the soluble yield of IMAC-refolded HBx was the incubation period, where extended incubation time was detrimental due to the probable oxidation of the glutathione molecules in the refolding buffer catalysed by the presence of Ni^{+2} ions [25]. This observation is in agreement with our earlier findings, where we observed that a net reducing environment is crucial for the improved stability of the HBx protein [13]. As the physicochemical environment surrounding HBx in the column becomes more and more oxidising, the efficiency of disulfide re-shuffling may be reduced. For some proteins which need to form disulfide bonds before other secondary structures develop [26,27], a redox-optimised physicochemical environment will be critical. The importance of incubation period on protein concentration is also reflected in Fig. 4B, where it is the second most important parameter affecting the concentration of IMAC-refolded HBx. The results in Table 2 also show that extended incubation time negatively impacts the two CQA parameters. Therefore the optimum on-column incubation time for the HBx protein was determined to be one day.

We next investigated the effects of different protein loads and buffer exchange gradients on HBx soluble yield and concentration

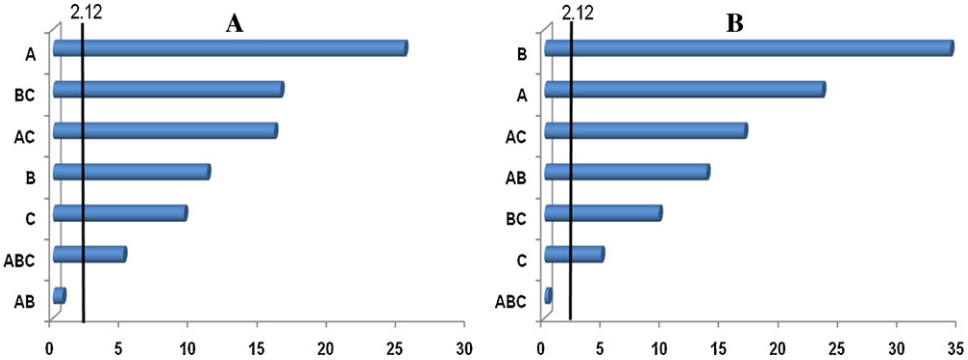

Fig. 4. Pareto plot showing the influence of the parameters (in order of their importance) on (A) HBx soluble yield and (B) concentration of the eluted HBx protein (under refolding conditions). (A) Incubation period (days); (B) protein loaded on to the column (mg); (C) rate of buffer exchange (CVs). Bars crossing the 2.12 line are significant at 95% confidence level.

Table 3
ANOVA table for (A) Eq. (4), and (B) Eq. (5).

	A						B					
Source	DF	Seq SS	Adj SS	Adj MS	F	P	DF	Seq SS	Adj SS	Adj MS	F	P
Main effects	3	6227.5	6227.46	2075.82	295.61	0	3	0.64665	0.64665	0.21555	613.7	0
2-Way interactions	2	3817.1	3817.08	1908.54	271.79	0	3	0.20918	0.20918	0.06973	198.52	0
3-Way interactions	1	187	187.04	187.04	26.64	0	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Residual error	17	119.4	119.38	7.02			17	0.00597	0.00597	0.00035		
Lack of fit	1	3.4	3.38	3.38	0.47	0.505	1	3.7E−05	3.7E−05	3.7E−05	0.1	0.755
Pure error	16	116	116	7.25			16	0.00593	0.00593	0.00037		
Total	23	10351					23	0.8618				

DF: degrees of freedom; Seq SS: sequential sums of squares; Adj SS: adjusted sum of squares; Adj MS: adjusted mean squares.

at a fixed on-column incubation period of 1 day in the refolding buffer. The results based on our prediction models (Eqs. (4) and (5)) are shown in the form of contour plots in Fig. 5. We observe that although the best soluble yields can be expected at lower protein loads (5 mg) and by using 5 CV buffer-exchange gradients (Fig. 5A), the protein concentration obtained would be relatively low (0.3–0.4 mg/ml). It is also clear from Fig. 5 that at a fixed protein load, soluble yields can be improved by increasing the length of buffer exchange gradients. The refolding kinetics of different proteins have been found to be significantly affected by the slope of the free energy landscape which in turn is directly influenced by the refolding physicochemical environment [28–31]. Our results demonstrate the importance of optimising the topography of the free energy landscape in on-column refolding studies, where the protein remains largely flexible to refold as only the 6His-containing portion of the protein remains bound to the charged Ni^{2+} resins. Therefore, the rate of change in the physicochemical environment will have a similarly important effect on governing the refolding behaviour such as efficiency and kinetics of the bound HBx proteins as it would have on free HBx proteins in solution.

Considering that a soluble yield of 60–70% at a concentration of >0.6 mg/ml would provide an acceptable product amount that enables their direct use for structural and other characterization studies, the following parameters, i.e. 10 mg protein load with 5 CV refolding buffer exchange and 1 day protein incubation, would provide the optimal parameter combination for IMAC refolding of the HBx protein. We set out to perform the IMAC refolding studies using those conditions and obtained a HBx soluble yield of 63% and eluted HBx concentration of 0.64 mg/ml, which validated our predicted models. HBx soluble yield and concentration were analysed by a chip-based electrophoresis method (Fig. 6). The bioactivity of the refolded HBx fraction was determined by ELISA to give a HBx refolding yield of 54 ± 12%. Although an iterative refolding strategy can be employed to refold the recovered HBx proteins in the stripping step to further improve process yield [20], the eluted protein concentration is likely to be lower than the allowable range specified in our QCA for this study, and this recycle step was therefore not considered.

As stated earlier, in our previous dilution refolding studies [14], we observed that the HBx protein has a slow refolding kinetics, where maximal yields (~55%) were achieved at a protein concentration of 0.1 mg/ml only after 3 days of incubation. Refolding the HBx protein using our optimised IMAC refolding process also achieved similar refolding yield but at an incubation period of one day and a refolding protein concentration of 10 mg/ml. Spatial isolation of the HBx molecules, achieved through protein immobilisation on an affinity chromatography based platform, seems to be vital for improving refolding productivity of the HBx protein. To compare the refolding performance of our IMAC refolding strategy with the conventional dilution refolding method, HBx was refolded and purified using two process schemes in parallel (Fig. 1), starting with a fixed protein load of 10 mg solublised denatured-reduced protein. The results show that the IMAC refolding process achieved a bioactive HBx amount and productivity that were 2.6 and 4.4-fold higher than the dilution refolding based process, respectively (Table 4). These results are not unexpected, considering the fact that the protein concentration range at which the dilution refolding is performed (i.e. ≥5 mg/ml) would almost certainly readily induce intermolecular protein interaction, thereby accelerating non-native off-pathway protein aggregation reactions. HBx protein production was thus found to be significantly improved through our optimised IMAC refolding process. From the results of this study, the effectiveness of employing a DoE methodology to generate process parameter information to guide

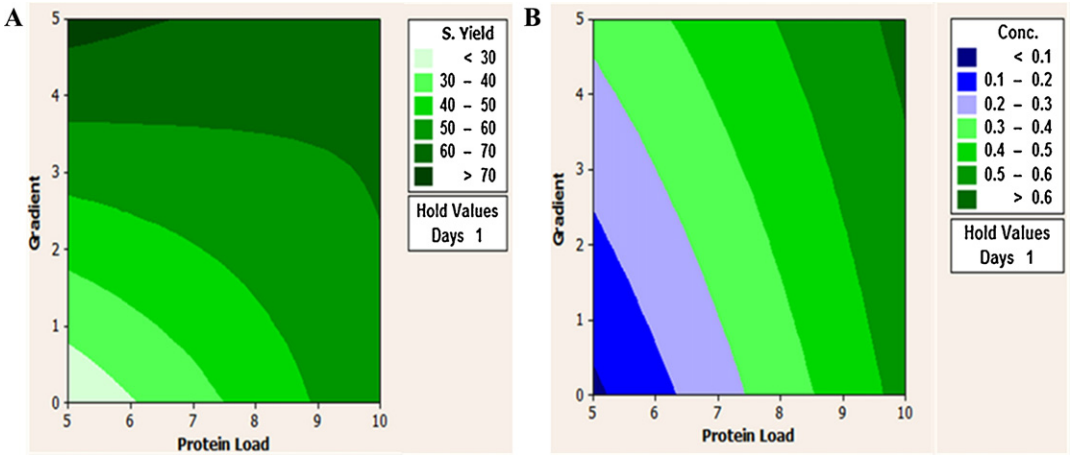

Fig. 5. (A) Contour plot of soluble yield (S. Yield) versus buffer exchange gradient and protein load at 1 day of incubation. (B) Contour plot of product concentration (Conc.) versus buffer exchange gradient and protein load at 1 day of incubation.

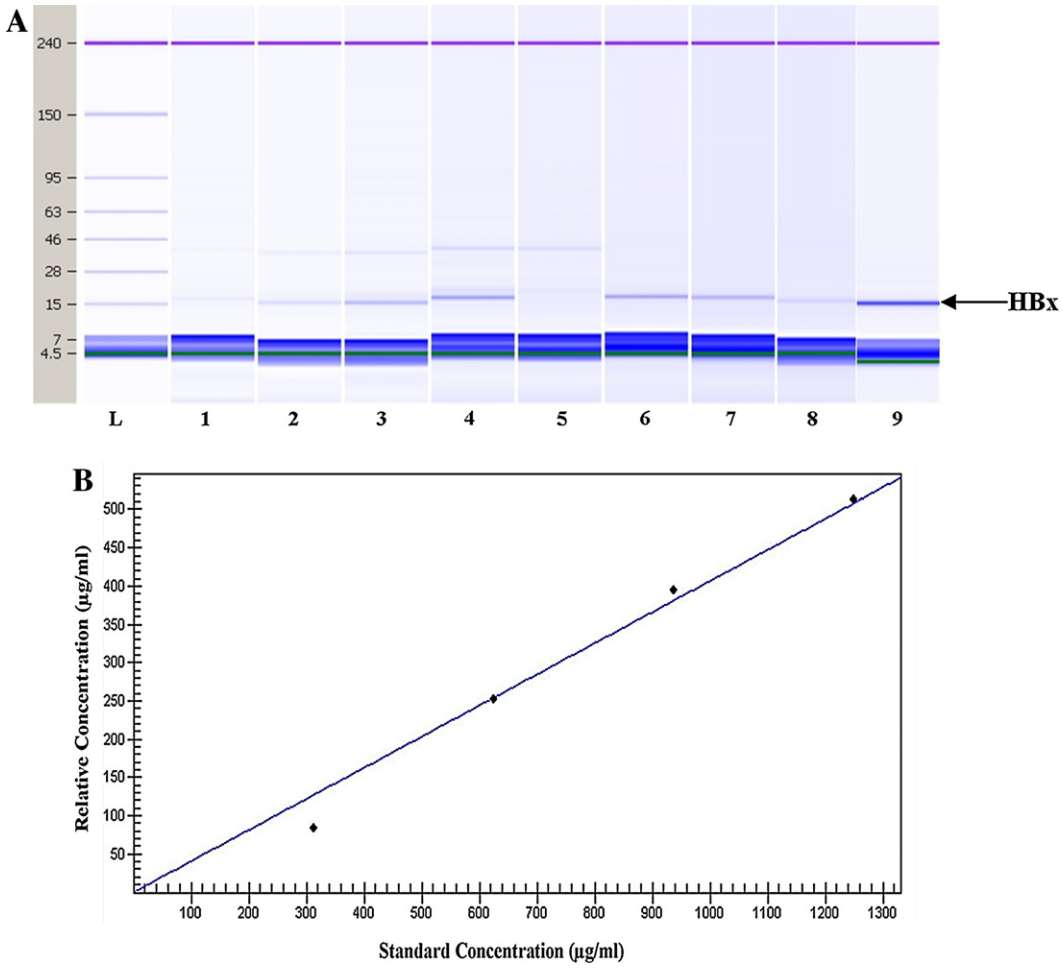

Fig. 6. (A) Electrophoresisresults obtained from Bioanalyser analysis of the optimised IMAC and dilution refolding processes. L: Protein marker; lanes 1–4: denatured reduced HBx IBs, diluted at different concentrations (consisting of 25, 50, 75, and 100%, respectively of the solubilised HBx IB solution); lane 5: flow through fractions; lanes 6–7: eluted fractions after IMAC refolding; 8: column stripping protein fractions; 9: dilution-refolded HBx. (B) Calibration curve obtained from denatured-reduced HBx samples in lanes 1–4 [$Y = 0.41^{*}X$; $R^2 = 0.9797$].

Table 4
Comparison of the HBx product characteristics obtained from process Schemes 1 and 2 at a fixed denatured-reduced protein load of 10 mg.

	Process	
	1	2
HBx purity (%)	>95	>95
HBx soluble yield (%)	63 ± 3	40 ± 4
HBx refolding yield (%)	54 ± 12	21 ± 3
HBx protein concentration (mg/ml)	>0.6	~2.0
Overall bioactive HBx (mg)	5.4 ± 1.2	2.1 ± 0.3
Refolding process productivity (mg/ml/h)	0.21 ± 0.05	0.05 ± 0.01

the design, development and optimisation of the HBx refolding process is clearly demonstrated.

4. Conclusions

Significant efforts are being directed to establish proper design spaces for the manufacturing of biopharmaceuticals which forms a crucial part in establishing QbD approaches, a mandate for current pharmaceutical manufacturing practices to meet stringent regulatory requirements. The development and optimisation of a bioprocess platform based on a DoE design methodology reported in this study is beneficial in establishing the optimised bioprocess right at the development stage, which saves time and resources while maximising information output. As shown in this study, the DoE approach provided useful information on the important parameters affecting the chromatography refolding behaviour for the HBx protein, while simultaneously providing the design space needed for subsequent regulatory approval procedures in the near future. The systematic methodology adopted for the current study is the first study of its kind for refolding the HBx protein, leading to the first intensified chromatography-based refolding platform for HBx production, to the best of our knowledge. The models developed through this study can be used to facilitate scale-up of the HBx IMAC refolding process which will then open the way for an efficient and cost-effective supply of the HBx protein to the scientific community striving to develop new drug candidates against HCC.

References

[1] D. Lavanchy, J. Viral Hepat. 11 (2004) 97.
[2] D.M. Parkin, P. Pisani, J. Ferlay, Int. J. Cancer 80 (1999) 827.
[3] A.S. Lok, Gastroenterology 132 (2007) 1586.
[4] M. Nassal, Future Virol. 4 (2009) 55.
[5] S. Benhenda, D. Cougot, M.A. Buendia, C. Neuveut, Adv. Cancer Res. (2009) 75.
[6] L.L. David, H. Nguyen, J. Hu, J. Cell. Physiol. 216 (2008) 289.
[7] H. Tang, N. Oishi, S. Kaneko, S. Murakami, Cancer Sci. 97 (2006) 977.
[8] M. Li, Z.-G. Su, J.-C. Janson, Protein Expr. Purif. 33 (2004) 1.
[9] Y. Chen, S.S.J. Leong, J. Chromatogr. A 1216 (2009) 4877.
[10] S.S.J. Leong, W.N. Chen, Chem. Eng. Sci. 63 (2008) 1401.
[11] A.S. Rathore, H. Winkle, Nat. Biotechnol. 27 (2009) 26.
[12] J. Harms, X. Wang, T. Kim, X. Yang, A.S. Rathore, Biotechnol. Prog. 24 (2008) 655.

[13] A. Basu, W. Chen, S. Leong, Appl. Microbiol. Biotechnol. 90 (2011) 181.
[14] A. Basu, W.N. Chen, S.S.J. Leong, Appl. Microbiol. Biotechnol. 90 (2011) 181.
[15] C. Jiang, L. Flansburg, S. Ghose, P. Jorjorian, A.A. Shukla, Biotechnol. Bioeng. 107 (2010) 985.
[16] A.A. Shukla, L. Sorge, J. Boldman, S. Waugh, Biotechnol. Appl. Biochem. 34 (2001) 71.
[17] X. Li, S.S.J. Leong, J. Chromatogr. A 1218 (2011) 3654.
[18] M.H. Hutchinson, G. Morreale, A.P.J. Middelberg, H.A. Chase, Biotechnol. Bioeng. 95 (2006) 724.
[19] M.H. Hutchinson, H.A. Chase, Biotechnol. Bioeng. 94 (2006) 1089.
[20] M.H. Hutchinson, H.A. Chase, J. Chromatogr. A 1128 (2006) 125.
[21] A. Basu, X. Li, S. Leong, Appl. Microbiol. Biotechnol. 92 (2011) 241.
[22] M. Langenhof, S.S.J. Leong, L.K. Pattenden, A.P.J. Middelberg, J. Chromatogr. A 1069 (2005) 195.
[23] Y. Chen, S.S.J. Leong, Process Biochem. 45 (2010) 1570.
[24] E.J. Freydell, L. van der Wielen, M. Eppink, M. Ottens, J. Chromatogr. A 1217 (2010) 7265.
[25] X. Geng, C. Wang, J. Chromatogr. B 849 (2007) 69.
[26] K. Tsumoto, T. Arakawa, L. Chen, Curr. Pharm. Biotechnol. 11 (2010) 285.
[27] S.S.J. Leong, A.P.J. Middelberg, Food Bioprod. Process. 84 (2006) 9.
[28] C.M. Dobson, Nature 426 (2003) 884.
[29] S.E. Radford, Trends Biochem. Sci. 25 (2000) 611.
[30] C.M. Dobson, Semin. Cell Devel. Biol. 15 (2004) 3.
[31] M. Oliveberg, Y.-J. Tan, M. Silow, A.R. Fersht, J. Mol. Biol. 277 (1998) 933.