

Titer estimation for quality control (TEQC) method: a practical approach for optimal production of protein complexes using the baculovirus expression vector system

Tsuyoshi Imasaki[#], Sabine Wenzel, Kentaro Yamada[#], Megan L. Bryant, Yuichiro Takagi^{*}

Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, Indiana 46202, USA

* Corresponding author

E-mail: ytakagi@iu.edu

[#]Present address: RIKEN Center for Life Science Technologies, Tsurumi-Ku, Yokohama, 230-0045, Japan (T.I.); Department of Pediatrics, Indiana University School of Medicine, 1044 W. Walnut Street, Indianapolis, Indiana 46202, USA (K.Y.)

Abstract

The baculovirus expression vector system (BEVS) is becoming the method of choice for expression of many eukaryotic proteins and protein complexes for biochemical, structural and pharmaceutical studies. Significant technological advancement has made generation of recombinant baculoviruses easy, efficient and user-friendly. However, there is a tremendous variability in the amount of proteins made using the BEVS, including different batches of virus made to express the same proteins. Yet, what influences the overall production of proteins or protein complexes remains largely unclear. Many downstream applications, particularly protein structure determination, require purification of large quantities of proteins in a repetitive manner, calling for a reliable experimental set-up to obtain the protein or protein complexes of interest consistently. During our investigation of optimizing the expression of the Mediator Head module, we discovered that the ‘initial infectivity’ was an excellent indicator of overall production of protein complexes. Further, we show that this initial infectivity can be mathematically described as a function of multiplicity of infection (MOI), correlating recombinant protein yield and virus titer. All these findings led us to develop the TiterEstimation for Quality Control (TEQC) method, which enables researchers to estimate initial infectivity, titer/MOI values in a simple and affordable way, and to use these values to quantitatively optimize protein expressions utilizing BEVS in a highly reproducible fashion.

Introduction

The baculovirus expression vector system (BEVS), introduced about 30 years ago [1-5], has become an essential tool for expression of many eukaryotic proteins and protein complexes in

insect cells [6-8]. BEVS has been proven powerful for structure determination of membrane proteins such as GPCRs [9], AMPA receptor [10], NMDA receptor [11], and channel rhodopsin [12], which are inherently difficult to work with, as well as cell surface receptors [13-15]. Moreover, development of the MultiBac system, which is suited for expression of multi-protein complexes, has opened new avenues for structural studies of multi-protein complexes [16-21]. Key technological advancements of the BEVS lie in the ease and efficient generation of recombinant baculoviruses harboring single genes [1, 5], or in case of the MultiBac system, harboring multiple genes, enabling expression of numerous proteins simultaneously [22-24]. Well-established standard procedures for BEVS allow a streamlined workflow, accounting for its ease of use and high adaptability to routine laboratory operations for many users. Despite of all these technological advances to simplify the generation of recombinant baculoviruses, when it comes to protein expression, the key indicator for optimal expression of proteins remains elusive. For example, there is no clear answer to the question of whether or not a user would obtain twice as much protein if twice as much baculovirus is added to the insect cells. The situation is a marked difference from the bacterial expression system in which key factors for optimization of protein expression (e.g. IPTG concentration) have been well established [25].

A lack of clarity regarding optimization of protein expression in BEVS may be attributed to conflicting results from previous studies [26-32]. In these studies, MOI (Multiplicity Of Infection: how many infectious units per cell) was considered to be a key indicator for expression of a model protein, β -galactosidase. While some of these studies showed that varying MOI values had no significant effects [26-28], others showed high MOI [29, 30] or even low MOI [28, 29, 31, 32] conditions resulted in optimal protein expression. The results of these studies did not provide a clear answer about the correlation between MOI and protein

production. The problem is a lack of understanding the physiological outcome of MOI values: for example, MOI=2 means having twice as much virus per cell, but how it affects the cells and how it represents protein expression levels are uncertain. Even if MOI is a good indicator, classical titer/MOI measurement methods, such as the plaque assay [33], and the immunohistochemical staining method (Clontech Inc), are too cumbersome and time-consuming to be practically used for high throughput expression studies. Other titer/MOI measurement methods such as qPCR [34], flow cytometric assays (FCA) [35], measurement of cell diameter change of post infected cells [36], and a cell line containing eGFP for titer measurement [37] are less tedious. However, they require specialized equipment, and setting them up is not trivial. Therefore, these methods are not practical for many users, whose goal is to produce proteins of their interest and not to measure MOI per se. Furthermore, the condition at which the titer is measured differs vastly from that at which protein expression is carried out, prompting a question as to how relevant such a measurement is for protein production. Taken all together, a simple and affordable experimental method to quantify and optimize expressions of proteins or protein complexes using BEVS needs to be established for many end users.

Mediator Head module is an essential sub-complex of the Mediator complex, which plays a key role in transcription regulation in eukaryotes [38]. We used BEVS to generate the Mediator Head module for structural and functional studies [16, 39]. We often encountered the problem of inconsistency in expression level of the Head module, which prompted us to investigate a way to optimize expression of the Mediator Head module. Our research led us to discover the ‘initial infectivity’ as an excellent indicator of overall production of protein complexes, and we devised a simple and affordable method to estimate this initial infectivity without the need for any specialized equipment or setting. Our finding led to the development of the ‘Titer Estimation for

Quality Control' (TEQC) method, which enables users to quantify and optimize expression of protein complexes in insect cells, ensuring quality control of protein expression in the BEVS.

Materials and Methods

Maintenance of insect cells

Hi5 and Sf9 cells were obtained from Expression Systems (Davis, CA). The cells were maintained in ESF921 media (Expression systems) in shaker flasks at 27°C. Both cells were split to a density of 0.5×10^6 cells/ml every 3 days.

Construction of baculovirus transfer vectors for expression of multi-protein complexes

All vectors used in this study are summarized in Supplemental Table S1. Construction of the baculovirus transfer vector for expression of modified Mediator Head module used for our structure determination was described in [16]. Construction of the baculovirus transfer vector bearing the genes encoding the yeast Mediator Middle module is as follows: Open reading frames (ORFs) of genes encoding Med7, Med21, Med9, Med4, Med10, Med19 and Med1, were amplified by PCR from the yeast (*S. cerevisiae*) genomic DNA. Med31 with a C-terminal HA and 10xhistidine tag (Med31-HA-10His) was amplified by PCR from pBacPAK9-Med31-HA-10His vector, which was generated by sub-cloning MED31 gene into BamHI and XhoI sites of pBacPAK9 followed by addition of HA and 10xHis tag sequences by two rounds of QuikChange site-directed mutagenesis (Agilent Technologies). The PCR products were cloned into SphI and SmaI sites (MCS1), and BamHI and HindIII

sites (MCS2) of the pFL vector using the SLIC method [40] in the following pair-wise combinations: Med7-Med21, Med9-Med4, Med10-Med31-HA-10xHis-tag, and pFL-Med19-Med1, yielding pFL-Med7-Med21 (pYT53), pFL-Med9-Med4 (pYT56), pFL-Med10-Med31-HA-10His (pYT55), and pFL-Med19-Med1 (pYT57), respectively. PmeI and AvrII fragment from pYT53 or pYT55 was sub-cloned into SpeI and NruI sites of pYT56 or pYT57, yielding pFL-Med4-Med7-Med21-Med9 (pYT87), and pFL-Med10-Med19-Med1-Med31HAHis (pYT85). Finally, PmeI and AvrII fragment of pYT87 was sub-cloned into SpeI and NruI vector of pYT85, yielding pFL-Med10-Med4-Med7-Med21-Med9-Med19-Med1-Med31HAHis (pYT90).

Construction of the baculovirus transfer vector bearing the genes encoding the subunits of yeast TFIIF, Tfg1, Tfg2 and Tfg3 is as follows: The C-terminal FLAG and TAP tagged TFG1 gene, yTFG1-FLAG-TAP, was amplified by PCR from pDt/glg2 vector [41] (gift from Ponticelli at State University of New York). ORFs of genes encoding yeast Tfg2 and Tfg3, were amplified by PCR from the yeast genomic DNA. The PCR products were cloned into SphI and SmaI sites (MCS1), and BamHI and HindIII sites (MCS2) of the pFL vector using the SLIC method [40], yielding pFL-Tfg1-TAP-Tfg2 (TI150). The PCR product for Tfg3 was cloned into BamHI and HindIII sites of pUCDM vector, yielding pUCDM-Tfg3 (pYT154). Fusion of the pTI150 and pYT154 vectors by Cre recombinase (NEB) yielded a transfer vector (TI153), containing all 3 genes encoding the yeast TFIIF. Construction of the baculovirus transfer vector bearing the genes encoding the subunits of yeast core TFIIF, is as follows: ORFs of genes encoding Tfb5, Rad3, Tfb2, Ssl1, Tbf4 and Tfb1-10xHis tagged, were amplified by PCR from pBacPAK9 vectors bearing corresponding genes as described previously [42]. The PCR products were sub-cloned into

SphI and SmaI sites (MCS1), and BamHI and HindIII sites (MCS2) of pFL, pUCDM, and pSPL vectors in pair-wise fashion, yielding pFL-Tfb5-Rad3 (TI132), pSPL-Tfb2-Ssl1 (pYT444), and pUCDM-Tfb1-10His-Tfb4 (pYT465), respectively. Fusion of all three vectors by Cre recombinase (NEB) yielded a transfer vector (TI419), containing all 5 genes encoding the yeast core TFIIF. Open reading frames (ORFs) of genes encoding yeast Cyclin C (CycC), and CDK8 genes were amplified from the yeast genomic DNA. DNA sequence encoding 10xhistidine tag was added to the C-terminus of CycC gene during PCR. The PCR products were cloned into SphI and SmaI sites (MCS1), and BamHI and HindIII sites (MCS2) of the pUCDM vector using the SLIC method [40], yielding pUCDM-CycC-10His-CDK8 (pYT379). Fusion of pYT379 vector and an empty pFL vector by Cre recombinase (NEB) yielded a transfer vector (TI600) for expression of CDK8-CycC-10xHis complex. Construction of human Taf8-Taf10 expression vector has been previously described [43].

Virus production and storage

Productions of recombinant baculoviruses in Sf9 cells were performed as described [23]. Liquid viruses were stored at 4°C. Frozen viruses were generated as described (Supplemental protocol) [44]. Frozen virus stocks were stored under liquid nitrogen.

Estimation of Multiplicity of Infection (eMOI)

The initial infectivity is defined as the infectivity 24 hours after infection (I24). An estimation of the initial infectivity is denoted as eI24. The multiplicity of infection (MOI),

which is a ratio of infectious units per cell, can be estimated using the following two equations:

$$eI_{24} = 1 - \left(\frac{C_1 - C_0}{(\alpha - 1)C_0} \right)$$

Estimation of MOI = eMOI

$$eMOI = -\ln(1 - eI_{24}) = -\ln \left(\frac{C_1 - C_0}{(\alpha - 1)C_0} \right)$$

Where α is a growth constant, α , is the growth rate of uninfected insect cells in a 24 hours period, C_0 is an initial cell number, and C_1 is the cell number 24 hours after addition of virus. The detailed procedure for estimation of infectivity and MOI is described in the supplementary protocol in Supplemental Information.

Titer measurement by Expression Systems, Inc

Our liquid viruses were shipped to Expression Systems (Davis, CA) for titer measurement. Titer measurements were performed via flow cytometric analysis (FCA) for expression of the baculoviral coat protein gp64 by Expression Systems, Inc as described [35]. The numbers of gp64 positive cells identified by FCA were used to calculate the infectivity of the baculovirus expressing the Mediator Head module.

Expression of the multi-protein complexes in the insect cells

200 ml culture of Hi5 cells (1.0×10^6 cells/ml) or Sf9 cells (1.5×10^6 cells/ml) was infected with the amount of viruses indicated. Infected cells were incubated at 27°C for 96

hours, and were harvested by centrifugation, frozen in liquid nitrogen, and stored as a pellet at -80°C until future use.

Purification of protein complexes

The Mediator Head module was purified as described previously [45]. The yeast Mediator Middle module, core TFIIF, CycC-CDK8, and human Taf8-Taf10, complexes were purified following the same protocol as well. Briefly, the cells from 200 ml culture were lysed in 50 ml of lysis buffer: 50 mM Hepes-KOH pH7.6, 400 mM potassium chloride, 10% Glycerol, and 5 mM β -mercaptoethanol, and 0.5 ml of 100 x protease inhibitor mix (6 mM leupeptin, 0.2 mM pepstatin A, 20 mM benzamidine, and 10 mM PMSF). Cell lysate was stirred for 30 min at 4°C followed by centrifugation at 100,000 g for 30 min at 4°C. The supernatant was loaded onto 0.5 ml of Ni-NTA (HIS-Select, Sigma-Aldrich). The resin was washed with 20 ml of the high salt buffer (50 mM Hepes-KOH pH7.6, 1 M potassium chloride, 10% Glycerol, 20 mM Imidazole, and 5 mM β -mercaptoethanol) followed by 5 ml of the low salt buffer (50 mM Hepes-KOH pH7.6, 100 mM potassium acetate, 10% Glycerol, 20 mM Imidazole, and 5 mM β -mercaptoethanol). Protein complexes were eluted from Ni-resin with the elution buffer (50 mM Hepes-KOH pH7.6, 100 mM potassium acetate, 10% Glycerol, 300 mM Imidazole, and 5 mM β -mercaptoethanol). The recombinant yeast TFIIF complex was purified via an IgG affinity column. Lysate from 200 ml culture prepared as described above, was loaded onto 0.5 ml of IgG resin (GE Healthcare). After wash with the high salt and low salt buffer, the recombinant yeast TFIIF was recovered by TEV protease digestion for 1 hour at room temperature. Concentrations of purified protein complexes were measured by Bradford assay.

Assay for solubility of the Head module subunits by quantitative western blotting

Cells infected with the various eMOIs were expressed in 50 ml culture of Hi5 cells (1.0×10^6 cells/ml), following the same experimental procedure described above. Cells were harvested in 1.5 ml tubes, cell pellets were frozen in liquid nitrogen, and stored at -80°C until use. Cell pellets were thawed and were resuspended in 100 μl lysis buffer. Lysates were centrifuged at 15,000 rpm for 10 minutes, at 4°C : supernatant and pellet were separated. Supernatants were mixed with 30 μl of 4x NuPAGE loading buffer (Life Technologies). Pellets were mixed with 120 μl of $1.5 \times$ NuPAGE loading buffer and solubilized by brief sonication. Samples were subjected to 4-12% NuPAGE (Invitrogen), transferred to PROTRAN membranes (Schleicher & Schuell), and probed for the Head module with anti-His tag monoclonal antibody (Thermo Scientific Pierce) for 10xHis-tagged Med17, anti-Med18 (anti-Srb5), and anti-Med11[46]. Detection was carried out using Dylight 680 goat anti-rabbit IgG (Thermo Scientific Pierce) for Med18 and Med11, and Dylight 800 goat anti-mouse IgG (Thermo Scientific Pierce) for 10-His tag on Med17 and scanning with an Odyssey infrared imaging system (LI-COR Biosciences). Bands were quantified using ImageJ software package [47]. 9 pmol (2 μg) of the purified Mediator Head module was used as standard.

Results and Discussion

The initial infectivity strongly correlates with overall protein complex production

In our previous work, we used the MultiBac baculovirus expression system as our method of

choice for high-yield, recombinant protein production [16]: As illustrated in Figure 1A, the multi-gene construct harboring all 7 genes encoding the Mediator Head subunits was generated, and integrated into baculovirus genome (DH10MultiBac) followed by virus production, expression and purification of the complex [16] (Figure 1A). We often encountered variability in production of the Mediator Head module when switching from one batch of baculovirus to the other, ranging from 3 mg per 1 L culture to 6 mg per 1 L culture (see below). As described in the introduction, given a lack of clarity regarding optimization of protein expression in BEVS, we were prompted to investigate what key factor, if any, is influencing expression of the Mediator Head module. We compared two different batches of baculoviruses harboring the genes for expression of the Mediator Head module: batch MB33 and batch MB88, which exhibited different expression levels: 4.3 mg per 1 L culture from batch MB33, and 6.4 mg per 1 L culture from batch MB88 (Figure 1B). We closely monitored both cell cultures every 24 hours by inspecting cell density and cell morphology. Typically, 24 hours after addition of baculovirus, there was already a marked difference between the two batches: as for batch MB88, the cell density was almost the same as the original cell density of 1.0×10^6 cells/ml, indicating growth arrest (Figure 1C). In addition, the majority of the cells showed rounded cell morphology, slightly increased overall size, as well as expansion of the nucleus, indicative of viral infection. The infectivity of batch MB88 at 24 hours appeared to be 100%. In contrast, after 24 hours the cells infected with batch MB33 virus had a cell density of 1.9×10^6 cells/ml, almost double of the initial cell density. Roughly 50% of cells appeared to be infected judged by their morphologies. The trend continued over the next time points, cells infected with batch MB88 showing 100 % infectivity after 24h, remained unchanged in their cell density until the end of the incubation. (Figure 1C). In contrast, infection with batch MB33 virus was incomplete after 24

hours, and uninfected cells continued to divide until 48 hours (Figure 1C), resulting in a noticeable difference in amount of the cells being harvested (illustrated in Figure 1B below). In short, the discrepancy in the overall expression levels of the Mediator Head module correlates with a difference in initial infectivity of the two batches of virus. Therefore, we hypothesize that the initial infectivity of the baculovirus, which we define here as infectivity 24 hours after virus addition: I_{24} , could be an excellent indicator of an overall production of the Mediator Head module.

To test our hypothesis, we pursued to develop a new and simple approach to determine initial infectivity (I_{24}) in order to examine the correlation between initial infectivities and expression levels of the Mediator Head module in-depth. Technically, flow cytometric assays [35], or the measurement of cell diameter change of post infected cells using Vi-CELL XR Cell Counter (Beckman Coulter) [36], could be used to measure an initial infectivity (I_{24}). However, as mentioned above, our goal was to devise an easy and affordable experimental method without the use of specialized equipment and setting so that many end users can utilize it. We reasoned that an initial infectivity (I_{24}) could be “estimated” by measuring cell densities: C_0 (number of cells per ml) being the initial cell density and C_1 being the cell density at 24 hours after addition of virus, based on the assumption that (i) infected cells lead to immediate growth arrest [48], resulting in no cell division, and that (ii) uninfected cells divide with a growth constant, α in 24 hours: α indicates how much fold change of uninfected cell growth for a 24 hours period, and it is defined as a ratio of cell densities at the time t and at the time, $(t+24 \text{ hours})$: $\alpha = \text{cell density at } (t+24)/\text{cell density at } t$. For example, $\alpha=2$ means that cell number doubles in every 24 hours. We like to emphasize at this point, that we were not detecting the number of infected cells directly. Rather we attempted to estimate how many cells are infected based on the consequence of virus

infection (infected cells led to cell arrest). Thus, we designate this value as estimate of I_{24} : eI_{24} .

Based on the assumption, the cell density at 24 hours after addition of virus, C_1 can be described as a sum of (i) the number of infected cells ($C_0 \times eI_{24}$) and (ii) the number of the uninfected cells that were divided $[(C_0 - C_0 \times eI_{24}) \times \alpha]$: $C_1 = C_0 \times eI_{24} + (C_0 - C_0 \times eI_{24}) \times \alpha$. Therefore, eI_{24} can be described as follows:

$$I_{24} \text{ estimate}(eI_{24}) = 1 - \frac{C_1 - C_0}{C_0(\alpha - 1)} \quad \text{– Equation 1 (Eq. 1)}$$

We evaluated if our new method provides estimates of infectivity, that are in line with those measured by commercially available FCA [35]. In this assay, Sf9 cells were infected by serially 10-fold diluted virus stock, and after 24 hours, the number of viral protein gp64 expressing cells were identified by flow cytometry. The initial infectivity can be determined by calculating the ratio between the number of gp64 expressing cells and total number of cells. In parallel, we set up our measurements using Sf9 cells under equivalent conditions in terms of the ratio between virus amount and cell number. The cell densities were measured and the initial infectivity for each virus amount was calculated using our formula (Eq. 1). Clearly, both methods are in good agreement for assessing the initial infectivity at almost all dilution points (Figure 2A), pointing to validity of our newly developed method.

With this new simple method in hands, we set up a series of the insect cell cultures with different amounts of the baculovirus expressing the Mediator Head module. Using our formula (Eq. 1), we calculated eI_{24} for each culture and plotted those values against protein yield of the Head module purified from each culture (Figure 2B). In this experiment, infected cells were incubated for 96 hours (our default setting). As shown in Figure 2B, there is a clear correlation between estimated initial infectivity (eI_{24}), and protein amount of the Mediator Head module: the Mediator Head module yield increases nearly proportionately to an increase in the value of eI_{24} until it reaches

close to 100%. Furthermore, we examined the correlation between eI_{24} and the protein complex yield when infected cells were incubated for a shorter time (48 and 72 hours) (Figure 2C). The length of incubation period influenced total yield of the complex. However, it did not influence the overall pattern of correlation between eI_{24} and yield (Figure 2C). Taken together, these data supports our hypothesis that our estimated initial infectivity, eI_{24} , could be a good indicator of the overall Mediator Head module expression in insect cells. The significance of this finding is that despite the complexity of expressing seven proteins simultaneously, the overall protein yield could be influenced by a ‘single’ factor, an initial infectivity, which can be easily estimated by counting cell density using a hemocytometer or cell counter.

Initial infectivity (I_{24}) of a recombinant baculovirus can be mathematically described as a function of multiplicity of infection (MOI), connecting MOI or titer to overall protein complex production

We attempted to mathematically describe initial infectivity, I_{24} , as a function of MOI by applying the mathematical formula first presented by Ellis and Delbruck in 1939 [49]: A virus infection is a stochastic event and the probability that number (n) of virus particles infected a cell with a given multiplicity of infection (MOI) can be approximated using Poisson distribution [49],

$$P(n) = MOI^n \times \frac{e^{-MOI}}{n!}, \text{ where } P(n) \text{ is the probability that an insect cell will get infected by}$$

number (n) of infectious units - baculoviruses. This equation is particularly applicable to

our case since an initial infectivity (I_{24}) was determined by a primary infection and is not influenced by secondary or tertiary infections. Infectivity as the average percentage of the insect cells that will become infected after addition of baculovirus with a given MOI can be described as: Infectivity = $P(n > 0) = 1 - P(0)$ and $P(0) = e^{-MOI}$.

Therefore, initial infectivity (I_{24}) can be described as a function of multiplicity of infection (MOI) as follows:

$$I_{24} = 1 - e^{-MOI} \text{ -- Equation 2 (Eq. 2)}$$

Key findings of this equation are two fold: first, a relationship between I_{24} and MOI is nonlinear (Figure. 2D); and second, more importantly, this equation links MOI directly to overall level of protein expression. Realizing this mathematical relationship between MOI and overall protein complex expression level is key discovery in this work.

We assume that I_{24} estimate (eI_{24}) = $1 - e^{-MOI \text{ estimate}}$ -- Equation 3 (Eq. 3).

By combining equations 1 and 3, I_{24} estimate = eI_{24} , and MOI estimate ($eMOI$) can be described as follows:

$$eMOI = -\ln(1 - eI_{24}) = -\ln\left(\frac{C_1 - C_0}{(\alpha - 1)C_0}\right) \text{ -- Equation 4 (Eq. 4)}$$

Since $MOI = \text{titer} / \text{cell density} \times \text{cell culture volume} (V_c) \times \text{virus volume} (V_{\text{virus}})$, then titer estimate, $eTiter$, can be described as follows:

$$eTiter = \frac{eMOI \times C_0 \times V_{\text{cell}}}{V_{\text{virus}}} \text{ -- Equation 5 (Eq. 5), where } V_{\text{virus}} \text{ is a virus volume added to the insect cell culture.}$$

We empirically tested our mathematical formula by utilizing the batch MB33 and batch MB88 viruses, which resulted in different Head module expression levels when infected with an equal volume of virus. The eI_{24} and the corresponding $eMOI$ values of batch MB33

and batch MB88 were determined to be: eI_{24} (batch MB33) = 0.51 (51%), $eMOI$ (batch MB33) = 1.32, and I_{24} (Batch MB88) = 0.90 (90%), $eMOI$ (batch MB88) = 3.1, respectively. Numbers of batch MB33, and batch MB88 and their Head module yields are nicely superimposed on the plot as a function of eI_{24} (Figure. 2B), as well as that of $eMOI$ (Figure. 2E), indicating the validity of our mathematical formula.

Our results so far led us to hypothesize the following as illustrated in Figure 3: first, a single factor, the initial infectivity (I_{24}) is a good indicator of overall protein complex expression level as is MOI despite of complexity of multiple proteins being expressed; second, the relationship between I_{24} and MOI is nonlinear (Figure 2D); third, I_{24} could be estimated simply by measuring cell density 24 hours after addition of baculovirus with Eq. 1; fourth, baculovirus “titer” can be estimated using Eq. 5; and finally and most importantly, using eI_{24} or $eMOI$ value of a given virus harboring multiple genes, expression of a protein complex can be quantified and thus, optimized (Figure 3).

The cell density measurement based method provides estimations of titer values that are in line with those measured by other titer measurement method

Our findings so far were very encouraging in terms of developing a simple and affordable method to estimate viral titer, MOI and infectivity. It will allow us quality control of protein complex production using the BEVS in insect cells at very early stages of the experiment. Moreover, the results from our cell density measurement based method are in good agreement with those from FCA method for assessing the initial infectivity of the baculovirus expressing

the Mediator Head module. We further evaluated whether our new method can also provide titer estimates that are in line with those measured by one of the conventional titer measurement assays. To this end, we compared titer values obtained from our formula (Eq. 5) (Figure 3) to those from commercially available flow cytometric assay based titer measurements, which provides titer values, that are in good correlation with those obtained by plaque assay [35]. Titers of a total of twelve baculoviruses were measured by both methods and compared (Figure 4A and 4B). Six of these virus do not contain any particular gene and consist of the control virus with known titer [1.0×10^9 Infectious Unit (IU) /ml] purchased from Expression Systems, Inc (Davis, CA), baculovirus generated from bacmid alone (DH10MultiBac) (Supplemental Figure S1A) [23], eYFP containing bacmid [50] (Supplemental Figure S1B), bacmids fused with several empty MultiBac transfer vectors, one vector (pFL) labeled as Vec1, (Supplemental Figure S1C), two vectors (pFL+pUCDM) labeled as Vec2, (Supplemental Figure S1D), or three vectors (pFL+pUCDM+pSPL) labeled as Vec3, (Supplemental Figure S1E) (Figure 4A). The other six baculoviruses contain genes for the protein complexes, which include the Mediator Head module, yeast TFIIF, human Taf8-Taf10 heterodimer, yeast CycC-CDK8 complex, yeast core TFIIF, and the Mediator Middle module, respectively (Figure 4B). As described in the supplemental protocol, we derived eTiter values from four different data points of different virus volumes using linear regression: Each measurement was repeated three times and averaged (Supplemental Figures S2-3). Comparing both methods for the viruses expressing protein complexes, the obtained titer values are in good agreement for five out of six viruses (Figure 4B). The other viruses, which do not contain genes for recombinant protein expression, and the one encoding the yeast core TFIIF genes titer values measured using our formula (Eq. 5) are on average 2.1 times higher than those determined by the FCA method. In this regard, there are

several other studies published comparing different titer measurement methods. For example, a 2-3 fold difference between the end-point dilution assay, and qPCR based assay was observed [51]; the analysis of end point dilution using GFP and anti-DBP (DNA Binding Protein) based assay showed as much as a 15-fold difference between the titer values [52]. Compared to these previous results, clearly, our measurement results are in a reasonable and acceptable range. Therefore, we conclude that our formula provides good titer estimations and thus, can be extended in its application to aid in the robust expression of many proteins.

As suggested, titer values can be dependent upon medium and cell types [35]. To address this, we tested if there is a difference in eTiter values between Sf9 cells and Hi5 cells. Since the FCA titer measurement service is available in only Sf9 cells, but not in Hi5 cells, we used our method to measure eTiter values of all twelve viruses described above using Hi5 cells and compared them to Sf9 cells (Figure 4C, Supplemental Figures S4). Although eTiter values of the control virus are similar in both cell lines, eTiter values of the other eleven MultiBac based viruses measured in Hi5 cells are approximately twice as high than those measured in Sf9 cells.

The condition at which $eI_{24} \approx 100\%$ (eMOI=3 or greater), provides the optimal or near optimal expression levels for a series of protein complexes involved in eukaryotic transcription by RNA polymerase

II

We applied our newly developed method to several other recombinant protein complexes involved in eukaryotic RNA polymerase II (Pol II) mediated transcription to see if their expression levels could be quantified and optimized in a similar fashion. Eukaryotic transcription

by Pol II is driven by a series of multi-protein complexes including the TATA-box protein associated factors (TAFs), general transcription factors, TFIID, and TFIIF, Mediator, and CDK8 module [53-55]. Generation of recombinant forms of these complexes and their sub-complexes will be essential for structural and functional studies. A set of expression baculoviruses for human Taf8-Taf10 complex (2 subunits)[43], yeast core TFIID (6 subunits), yeast TFIIF (3 subunits), yeast CDK8-CycC (2 subunits), yeast Mediator Middle module (8 subunits), and the yeast Mediator Head module (7 subunits) were generated. Following our protocol (Supplemental Protocol), eI_{24} value of each virus was determined followed by calculating eMOI, and eTiter (Figure 5). These complexes were expressed in 8-9 different infection conditions ranging from 18% infectivity (eMOI=0.2) to ~100% (eMOI=3-5) (Figure 5). Each complex at each infection point was expressed in a 200 ml culture of Sf9 or Hi5 cells followed by affinity purification and quantification of the purified complex. The yield of each protein complex was plotted against eI_{24} . We also used the corresponding eMOI to display because eMOI is intuitively easy to understand: for instance, eMOI=2 indicates adding twice as much volume of virus as eMOI=1. There is a good correlation between the eI_{24} of the expression baculovirus and the overall protein complex production (Figure 5). Protein complex yield peaked at or near infection saturation point of eMOI being 3 or greater (Figure 5) except for the case of the Mediator Middle module: its yield peaked at an earlier infectivity point (39.3%; eMOI=0.5) (Figure 5A). However, the protein complex yield difference between eMOI=0.5 and eMOI=3-5 is relatively small with 10 mg/L and 9 mg/L, respectively, accounting for only a 10% difference. In other words, the eMOI=3-5 is nearly optimal. Key conclusion is that the condition at which eI_{24} =100% or $eMOI \geq 3$ provides the optimal, or near optimal expression levels for all cases tested. We call this the “eMOI=3 or greater” rule. It is not entirely surprising to see that optimal protein expression is

observed when cells are fully infected within the first 24 hours, considering a delay in infection due to lower virus load consequently postpones reaching the level of maximum amount of recombinant protein producing cells and ultimately yield in the given time-frame. Our “eMOI=3 or greater” captures such condition for an optimal protein production.

We further tested the idea that once the cells are more than 95 % infected ($\text{eMOI} \geq 3$), any additional excess virus would not necessarily increase overall expression level of protein complexes. To test this idea, the Mediator Head module, hTaf8-Taf10 and yeast core TFIID were chosen because the Head is our model complex and other two are on different ends of the spectrum in size and expression levels: hTaf8-Taf10 is considered a relatively small complex (81 kDa), and was highly expressed (Figure 5A) while yeast core TFIID is a relatively large complex (6 subunits, 320 kDa) and not expressed well (Figure 5B). For all three complexes (Supplemental Figure S5), additional excess virus did not improve overall yield of the protein complexes, suggesting that once the cells are fully infected, further addition of expression virus may not improve overall expression of protein complexes. During our study, we noticed a marked difference in expression levels of protein complexes between Hi5 and Sf9 cells. Performance of Hi5 cells exceeded Sf9 cells in 4 out of 6 complexes (Figure 5A, 5D-F). For expression of yeast core TFIID (Figure 5B) and TFIIF (Figure 5C), Sf9 cells performed better than Hi5 cells, suggesting that both cell lines should be tested to obtain maximum yield of protein complex of interest.

Finally, we tested if an increase in yield of protein complex could be attributed to an increase in overall expression of the complex or an increase in solubility of the subunits. We used the Mediator Head module as a test case since it showed most change in its yield between low and high eMOI (Figure 2D). We compared soluble and insoluble fractions of the Head module with

different eMOI ranging from eMOI=0.1 (10% infected initially) to eMOI=4 (~100% infected) by immunoblotting using antibodies against Med17, Med18 and Med11 subunits (Supplemental Figure S6A-C). The ratio between soluble and insoluble fractions of Med17, Med18 and Med11 appears to stay constant, suggesting that an increase of I_{24} or eMOI leads to an increase in overall production of the Mediator Head subunits.

The results of a total of six multi-protein complexes are consistent with our hypothesis that eI_{24} (or eMOI) can be an excellent indicator of expression levels of multi-protein complexes. By varying eI_{24} or eMOI value, the optimum expression points of all complexes can be identified. But in the end, the condition at which eI_{24} =100% or eMOI ≥ 3 provides the optimal, or near optimal expression levels (“eMOI=3 or greater” rule). In essence, our simple and affordable cell density measurement based method provides eI_{24} or eMOI value of a recombinant baculovirus. Researchers (end users) can use these values to quantitatively optimize expression levels of proteins or protein complexes of their interest. We named our newly developed quantitative method “TEQC” method: “Titer Estimation for Quality Control” (TEQC) of protein or protein complex production.

TEQC method can be utilized to achieve reproducible expression of protein complexes to ensure quality control

We tested whether we could use the eMOI value to achieve reproducible expression of protein complexes, resulting in a production of the recombinant protein complexes with maximum yield consistently. The conditions for maximum expression levels of 6 protein complexes are as follows: in the Hi5 cells, the Mediator Head module peaked at eMOI of 3.7; hTaf8-Taf10 at

eMOI of 3.5, yeast CDK8-CycC at eMOI = 3.5, and Mediator middle module at eMOI= 0.5; in SF9 cells, yeast core TFIH at eMOI of 3.0, and yeast TFIIF at eMOI of 3.0 (Figure 5), respectively. For simplicity, we are referring to eMOI only. All these six protein complexes were expressed using the same optimum eMOI and cell line but were expressed on different days. These independent expressions were repeated three times for all complexes. The yield of each purified protein complex was measured, the numbers were averaged, and compared to the original expression level that provided the optimum eMOI value in that cell line. For all tested complexes, the yield of purified protein was highly reproducible (Figure 6A).

One of the BEVS benefits is that once a virus is made, it is stable for a long period of time, ranging from up to couple of years if stored at 4 °C to much longer time spans for frozen stocks, allowing it be used for multiple rounds of infections and experiments. Over time though, the infectivity of the virus decreases. To address the situation in which one would like to use expression baculovirus that has declined titer due to prolonged storage at 4 °C we hypothesize as follows: as long as the eI_{24} or eMOI value is the same, an old virus with declined titer could maintain the same level of expression of protein complex as when it was newly made. To test this, we set up expressions of four protein complexes described above with eMOI from 1.0 to 4.0, using the old viruses that had been kept at 4 °C for an extended period of time (8 to 10 months). Their eTiters had been declined as follows: Mediator Head from 7.3×10^8 to 4.0×10^8 (IU/ml), hTaf8-Taf10 1.2×10^9 (IU/ml) from to 6.7×10^8 (IU/ml), core TFIH from 4.5×10^8 to 3.3×10^8 (IU/ml), and TFIIF from 5.1×10^8 to 3.8×10^8 (IU/ml), respectively. eMOI values ranging from 1 to 4 were achieved by increasing the volume of each virus to compensate for the loss of virus titer. The yield of each purified protein complex was measured and compared to the original expression level (Figure 6B-E, blue lines) by superimposing the data onto the plots of

the original expression data indicated by red lines (Figure 6B-E). Overall, when the eI₂₄ or eMOIs were adjusted, the expression level profiles of the protein complexes generated by the old viruses were similar to those generated by their original viruses (Figure 6B-E), strongly suggesting that eI₂₄ or eMOI is the key indicator of expression level of protein complexes.

Conclusion

This work has revealed the connection among initial infectivity, MOI, and an overall recombinant protein production. We developed a simple and affordable cell density measurement based method to estimate initial infectivity, MOI, and titer, leading toward the development of the TEQC method for optimization of a production of protein complexes using the BEVS. Key finding is that optimal or near optimal expression levels are reached when eMOI ≥ 3 , which our “eMOI=3 or greater” rule. Since expression of single proteins could be viewed as “one” subunit multi-protein complex, the TEQC method should be applicable to expression of single proteins as well. Additionally, expression of a protein complex in insect cells was carried out in this work by infecting a single virus harboring multiple genes encoding the subunits of the complex, and not by co-infecting multiple viruses. The strategy for optimization of protein complex expression using the co-infection method will be a future subject of investigation.

Author contributions

T. I. and Y. T. conceived the idea and designed experiments; T. I., M.B., K.Y., S. W., performed the experiments. S.W. T. I. and Y. T. analyzed the data; Y. T. supervised the study; and T. I., S.W., and Y.T. prepared figures and wrote the paper.

Acknowledgements

We thank Dr. B. Hayes for providing the raw FCA data and technical help. We thank Drs. H. Hundley, I. Berger and S. Trowitzsch for critical reading of the manuscript. We thank Drs. I. Berger and S. Trowitzsch for providing human Taf8-Taf10 construct and Dr. F. Ponticelli for providing pDt/g1g2 vector. We thank Dr. H. Noguchi for his help generating a part of the data. This research was supported by US National Science Foundation grant MCB-1157688, the National Institutes of Health (R01 GM111695), and Showalter Trust Fund to (Y.T.).

References:

1. van Oers MM. Opportunities and challenges for the baculovirus expression system. *J Invertebr Pathol.* 2011;107 Suppl:S3-15. doi: 10.1016/j.jip.2011.05.001. PubMed PMID: 21784228.
2. van Oers MM, Pijlman GP, Vlak JM. Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *The Journal of general virology.* 2015;96(Pt 1):6-23. doi: 10.1099/vir.0.067108-0. PubMed PMID: 25246703.
3. Miller LK. Baculoviruses as gene expression vectors. *Annual review of microbiology.* 1988;42:177-99. doi: 10.1146/annurev.mi.42.100188.001141. PubMed PMID: 3059993.
4. Luckow VA, Summers MD. Trends in the development of baculovirus expression vectors. *Nat Biotechnol.* 1988;6(1):47-55.
5. Smith GE, Fraser MJ, Summers MD. Molecular Engineering of the *Autographa californica* Nuclear Polyhedrosis Virus Genome: Deletion Mutations Within the Polyhedrin Gene. *J Virol.* 1983;46(2):584-93. PubMed PMID: 16789242; PubMed Central PMCID: PMC255161.
6. Kost TA, Condreay JP, Jarvis DL. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol.* 2005;23(5):567-75. Epub 2005/05/07. doi: 10.1038/nbt1095. PubMed PMID: 15877075; PubMed Central PMCID: PMC3610534.
7. Massotte D. G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. *Biochim Biophys Acta.* 2003;1610(1):77-89. PubMed PMID: 12586382.
8. Assenberg R, Wan PT, Geisse S, Mayr LM. Advances in recombinant protein expression for use in pharmaceutical research. *Curr Opin Struct Biol.* 2013;23(3):393-402. doi: 10.1016/j.sbi.2013.03.008. PubMed PMID: 23731801.
9. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature.* 2011;477(7366):549-55. Epub 2011/07/21. doi: 10.1038/nature10361. PubMed PMID: 21772288; PubMed Central PMCID: PMC3184188.
10. Sobolevsky AI, Rosconi MP, Gouaux E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature.* 2009;462(7274):745-56. doi: 10.1038/nature08624. PubMed PMID: 19946266; PubMed Central PMCID: PMC2861655.
11. Karakas E, Furukawa H. Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science.* 2014;344(6187):992-7. doi: 10.1126/science.1251915. PubMed PMID: 24876489; PubMed Central PMCID: PMC4113085.
12. Kato HE, Zhang F, Yizhar O, Ramakrishnan C, Nishizawa T, Hirata K, et al. Crystal structure of the channelrhodopsin light-gated cation channel. *Nature.* 2012;482(7385):369-74. doi: 10.1038/nature10870. PubMed PMID: 22266941; PubMed Central PMCID: PMC4160518.
13. Janda CY, Waghray D, Levin AM, Thomas C, Garcia KC. Structural basis of Wnt recognition by Frizzled. *Science.* 2012;337(6090):59-64. doi: 10.1126/science.1222879. PubMed PMID: 22653731; PubMed Central PMCID: PMC3577348.
14. Tanji H, Ohto U, Shibata T, Miyake K, Shimizu T. Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science.* 2013;339(6126):1426-9. doi: 10.1126/science.1229159. PubMed PMID: 23520111.

15. Ohto U, Shibata T, Tanji H, Ishida H, Krayukhina E, Uchiyama S, et al. Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature*. 2015;520(7549):702-5. doi: 10.1038/nature14138. PubMed PMID: 25686612.
16. Imasaki T, Calero G, Cai G, Tsai KL, Yamada K, Cardelli F, et al. Architecture of the Mediator head module. *Nature*. 2011;475(7355):240-3. Epub 2011/07/05. doi: 10.1038/nature10162. PubMed PMID: 21725323.
17. Schreiber A, Stengel F, Zhang Z, Enchev RI, Kong EH, Morris EP, et al. Structural basis for the subunit assembly of the anaphase-promoting complex. *Nature*. 2011;470(7333):227-32. doi: 10.1038/nature09756. PubMed PMID: 21307936.
18. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature*. 2014;516(7531):355-60. doi: 10.1038/nature14008. PubMed PMID: 25409142.
19. Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crepin T, et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature*. 2014;516(7531):361-6. doi: 10.1038/nature14009. PubMed PMID: 25409151.
20. Bleichert F, Botchan MR, Berger JM. Crystal structure of the eukaryotic origin recognition complex. *Nature*. 2015;519(7543):321-6. doi: 10.1038/nature14239. PubMed PMID: 25762138; PubMed Central PMCID: PMC4368505.
21. Barford D, Takagi Y, Schultz P, Berger I. Baculovirus expression: tackling the complexity challenge. *Curr Opin Struct Biol*. 2013;23(3):357-64. Epub 2013/05/01. doi: 10.1016/j.sbi.2013.03.009. PubMed PMID: 23628287.
22. Berger I, Fitzgerald DJ, Richmond TJ. Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol*. 2004;22(12):1583-7. PubMed PMID: 15568020.
23. Fitzgerald DJ, Berger P, Schaffitzel C, Yamada K, Richmond TJ, Berger I. Protein complex expression by using multigene baculoviral vectors. *Nat Methods*. 2006;3(12):1021-32. PubMed PMID: 17117155.
24. Bieniossek C, Imasaki T, Takagi Y, Berger I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci*. 2012;37(2):49-57. Epub 2011/12/14. doi: 10.1016/j.tibs.2011.10.005. PubMed PMID: 22154230.
25. Tolia NH, Joshua-Tor L. Strategies for protein coexpression in *Escherichia coli*. *Nat Methods*. 2006;3(1):55-64. doi: 10.1038/nmeth0106-55. PubMed PMID: 16369554.
26. Murhammer DW, Gooch CF. Scaleup of insect cell culture: protective effects of Pluronic F-68. *Nat Biotechnol*. 1988;6(12):1411-8.
27. Schopf B, Howaldt MW, Bailey JE. DNA distribution and respiratory activity of *Spodoptera frugiperda* populations infected with wild-type and recombinant *Autographa californica* nuclear polyhedrosis virus. *Journal of biotechnology*. 1990;15(1-2):169-85. PubMed PMID: 1367465.
28. Wong KT, Peter CH, Greenfield PF, Reid S, Nielsen LK. Low multiplicity infection of insect cells with a recombinant baculovirus: The cell yield concept. *Biotechnol Bioeng*. 1996;49(6):659-66. doi: 10.1002/(SICI)1097-0290(19960320)49:6<659::AID-BIT7>3.0.CO;2-N. PubMed PMID: 18626861.
29. Licari P, Bailey JE. Factors influencing recombinant protein yields in an insect cell-baculovirus expression system: multiplicity of infection and intracellular protein degradation. *Biotechnol Bioeng*. 1991;37(3):238-46. doi: 10.1002/bit.260370306. PubMed PMID: 18597361.

30. Neutra R, Levi BZ, Shoham Y. Optimization of protein-production by the baculovirus expression vector system in shake flasks. *Appl Microbiol Biotechnol*. 1992;37(1):74-8. PubMed PMID: 1368502.
31. Power JF, Reid S, Radford KM, Greenfield PF, Nielsen LK. Modeling and optimization of the baculovirus expression vector system in batch suspension culture. *Biotechnol Bioeng*. 1994;44(6):710-9. doi: 10.1002/bit.260440607. PubMed PMID: 18618831.
32. Radford KM, Cavegn C, Bertrand M, Bernard AR, Reid S, Greenfield PF. The indirect effects of multiplicity of infection on baculovirus expressed proteins in insect cells: secreted and non-secreted products. *Cytotechnology*. 1997;24(1):73-81. doi: 10.1023/A:1007962903435. PubMed PMID: 22358599; PubMed Central PMCID: PMC3449608.
33. Dulbecco R, Vogt M. Plaque formation and isolation of pure lines with poliomyelitis viruses. *J Exp Med*. 1954;99(2):167-82. PubMed PMID: 13130792; PubMed Central PMCID: PMC2180341.
34. Lo H-R, Chao Y-C. Rapid Titer Determination of Baculovirus by Quantitative Real-Time Polymerase Chain Reaction. *Biotechnol Progress*. 2008;20(1):354-60.
35. Mulvanian T, Hayes D. A Flow Cytometric Assay for Rapid, Accurate Determination of Baculovirus Titers. *BioProcessing J*. 2004;3:47-53.
36. Janakiraman V, Forrest WF, Seshagiri S. Estimation of baculovirus titer based on viable cell size. *Nature protocols*. 2006;1(5):2271-6. doi: 10.1038/nprot.2006.387. PubMed PMID: 17406467.
37. Hopkins R, Esposito D. A rapid method for titrating baculovirus stocks using the Sf-9 Easy Titer cell line. *Biotechniques*. 2009;47(3):785-8. doi: 10.2144/000113238. PubMed PMID: 19852765.
38. Kornberg RD. Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci*. 2005;30(5):235-9. PubMed PMID: 15896740.
39. Cai G, Imasaki T, Yamada K, Cardelli F, Takagi Y, Asturias FJ. Mediator head module structure and functional interactions. *Nature structural & molecular biology*. 2010;17(3):273-9. Epub 2010/02/16. doi: 10.1038/nsmb.1757. PubMed PMID: 20154708; PubMed Central PMCID: PMC2925518.
40. Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods*. 2007;4(3):251-6. PubMed PMID: 17293868.
41. Yang C, Khapersky DA, Hou M, Ponticelli AS. Improved methods for expression and purification of *Saccharomyces cerevisiae* TFIIF and TFIIF; identification of a functional *Escherichia coli* promoter and internal translation initiation within the N-terminal coding region of the TFIIF TFG1 subunit. *Protein Expr Purif*. 2010;70(2):172-8. doi: 10.1016/j.pep.2009.09.021. PubMed PMID: 19818408; PubMed Central PMCID: PMC2827679.
42. Takagi Y, Komori H, Chang WH, Hudmon A, Erdjument-Bromage H, Tempst P, et al. Revised subunit structure of yeast transcription factor IIH (TFIIH) and reconciliation with human TFIIH. *J Biol Chem*. 2003;278(45):43897-900. PubMed PMID: 14500720.
43. Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, et al. The architecture of human general transcription factor TFIID core complex. *Nature*. 2013;493(7434):699-702. doi: 10.1038/nature11791. PubMed PMID: 23292512.
44. Wasilko DJ, Lee SE, Stutzman-Engwall KJ, Reitz BA, Emmons TL, Mathis KJ, et al. The titerless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with

- recombinant baculovirus. *Protein Expr Purif.* 2009;65(2):122-32. doi: 10.1016/j.pep.2009.01.002. PubMed PMID: 19174191.
45. Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, et al. Head module control of mediator interactions. *Mol Cell.* 2006;23(3):355-64. Epub 2006/08/04. doi: S1097-2765(06)00412-6 [pii]
10.1016/j.molcel.2006.06.007. PubMed PMID: 16885025.
46. Takagi Y, Kornberg RD. Mediator as a general transcription factor. *J Biol Chem.* 2006;281(1):80-9. PubMed PMID: 16263706.
47. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012;9(7):671-5. PubMed PMID: 22930834.
48. Volkman LE, Summers MD, Hsieh CH. Occluded and nonoccluded nuclear polyhedrosis virus grown in *Trichoplusia ni*: comparative neutralization comparative infectivity, and in vitro growth studies. *J Virol.* 1976;19(3):820-32. PubMed PMID: 787558; PubMed Central PMCID: PMC354923.
49. Ellis EL, Delbruck M. The Growth of Bacteriophage. *J Gen Physiol.* 1939;22(3):365-84. PubMed PMID: 19873108; PubMed Central PMCID: PMC2141994.
50. Fitzgerald DJ, Schaffitzel C, Berger P, Wellinger R, Bieniossek C, Richmond TJ, et al. Multiprotein expression strategy for structural biology of eukaryotic complexes. *Structure.* 2007;15(3):275-9. Epub 2007/03/16. doi: S0969-2126(07)00072-X [pii]
10.1016/j.str.2007.01.016. PubMed PMID: 17355863.
51. Lo HR, Chao YC. Rapid titer determination of baculovirus by quantitative real-time polymerase chain reaction. *Biotechnology progress.* 2004;20(1):354-60. doi: 10.1021/bp034132i. PubMed PMID: 14763863.
52. Kwon MS, Dojima T, Toriyama M, Park EY. Development of an antibody-based assay for determination of baculovirus titers in 10 hours. *Biotechnology progress.* 2002;18(3):647-51. doi: 10.1021/bp020298s. PubMed PMID: 12052086.
53. Boeger H, Bushnell DA, Davis R, Griesenbeck J, Lorch Y, Strattan JS, et al. Structural basis of eukaryotic gene transcription. *FEBS Lett.* 2005;579(4):899-903. Epub 2005/02/01. doi: S0014-5793(04)01406-1 [pii]
10.1016/j.febslet.2004.11.027. PubMed PMID: 15680971.
54. Galbraith MD, Donner AJ, Espinosa JM. CDK8: a positive regulator of transcription. *Transcription.* 2010;1(1):4-12. doi: 10.4161/trns.1.1.12373. PubMed PMID: 21327159; PubMed Central PMCID: PMC3035184.
55. Egly JM, Coin F. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst).* 2011;10(7):714-21. doi: 10.1016/j.dnarep.2011.04.021. PubMed PMID: 21592869.

Figure Legends

Figure 1. Discrepancy in expression levels of the Mediator Head module was observed

when using two different batches of the recombinant baculoviruses, prompting our

investigation to seek a key indicator. (A) Schematic diagram of the workflow of the MultiBac

baculovirus expression vector system (BEVS) for expression of the Mediator Head module. The

single transfer vector harboring genes encoding subunits of the Mediator Head module was

integrated into a baculovirus gene followed by virus production, expression and purification of

the complex (top). SDS-PAGE of purified Mediator Head module is shown (bottom on right).

The structure using the recombinant Head module was determined by X-ray crystallography

(PDBID: 3RJ1) (bottom on left). Med17 (blue), Med11 (purple), Med22 (green), Med6 (yellow),

Med8 (red), Med18 (cyan), and Med 20 (orange) with electron density map (light pink) (B)

Quantity of purified Mediator Head Module varies notably for different batches of the expression

baculovirus (top). Batch MB33 in red and MB88 in dark blue (C) Time course for cell density

infected by batch MB33 and batch MB88 viruses during expression from time 0 to up to 96

hours.

Figure 2. Infectivity of insect cells 24 hours after addition of virus could serve as a good

indicator for the expression level of the Mediator Head module. (A) Comparison between the

initial infectivity estimates (eI_{24}) derived from the cell density measurement based method and

I_{24} (FCA) from flow cytometric assay (FCA). Sf9 cells were infected with 10-fold serially

diluted recombinant baculovirus expressing the Mediator Head module as indicted in Figure 2A.

24 hours after addition of the virus, infectivity of each dilution point was determined by our cell

density measurement based method as well as by FCA (Expression Systems Inc.) in which the

percentage of gp64 expressing cells was calculated. X-axis: Virus volume/ total cell number ($V_{\text{virus}}/\text{cell} \times 5.0 \times 10^{-5}$); Y-axis: initial infectivity (%). eI₂₄ are shown in dark blue and FCA I₂₄ are shown in green. (B) Initial infectivity (I₂₄) vs. yield of the Mediator Head Module. A total of 8 x 200 ml cultures were infected by various amounts of the baculovirus harboring the genes that encode the Mediator Head module. SDS-PAGE of purified Mediator Head module in eight different infectivity conditions is displayed on the right. The protein complex yield from each condition was measured and plotted against eight different infectivities on the left. The estimated infectivity was ranging from 45% to 98.1%. The cell culture conditions of batch MB33 and batch MB88 were superimposed onto Figure 2B indicated by red closed circles and red arrows. (C) Correlation between eI₂₄ and the protein complex yield when infected cells were incubated for 48, 72, and 96 hours. A total of 5 x 200 ml cultures were infected by various amounts of the baculovirus harboring the genes that encode the Mediator Head module, and incubated for shorter time periods (48 and 72 hours). The protein complex yield at each infectivity from the culture incubated for 48 or 72 hours was measured, and plotted against five different infectivities. The data from the culture incubated for 96 hours shown in Figure 2B, was superimposed in the same plot. (D) Relationship between infectivity and MOI. Non-linear nature of mathematical relationship between infectivity and multiplicity of infection (MOI) is displayed as table on the left and graph on the right. (E) eMOI vs. yield of the Mediator Head Module. Yield of the Mediator Head Module was plotted against eMOI, which was calculated from eI₂₄ values using Eq.4. The cell culture conditions of batch MB33, and batch MB88 were superimposed onto Figure 2E, indicated by red closed circles and red arrows.

Figure 3. Schematic diagram for our hypothesis illustrating the mathematical relationship among protein production, eI₂₄, eMOI, and eTiter The relationship among protein production, eI₂₄, eMOI, and eTiter was displayed as schematic diagram.

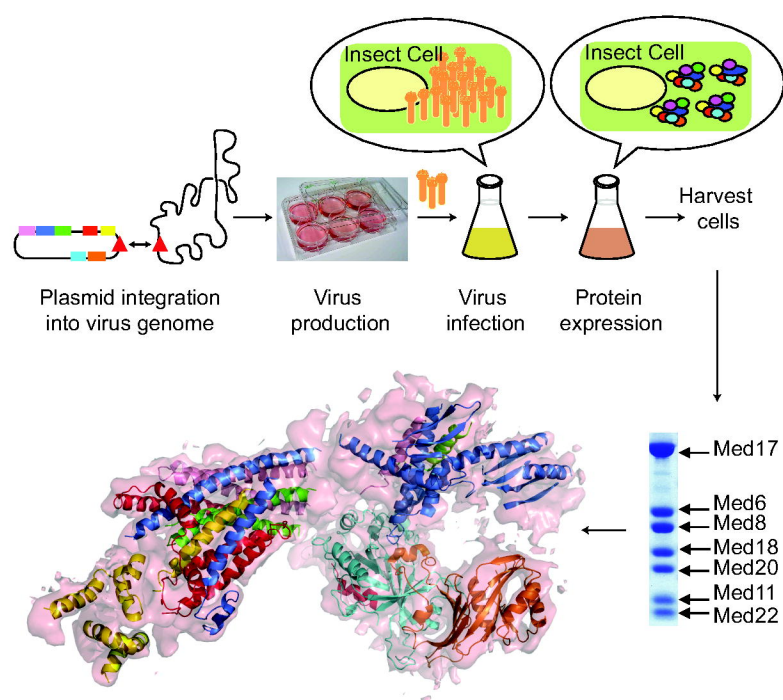
Figure 4. Comparison of titer values by the cell density measurement based method vs. commercially available FCA based titer measurement. (A) Titer values were obtained from the cell density measurement based method and commercially available flow cytometric assay (FCA) based titer measurement for a total of twelve baculoviruses, and these values were compared. All measurements were performed in Sf9 cells. Titers of the control virus with known titer (1.0×10^9 IU/ml) purchased from Expression Systems, baculoviruses generated from bacmid alone [23], eYFP containing bacmid [50], bacmids fused with several empty Multibac transfer vectors, respectively. Each measurement was repeated three times and averaged except for the control virus. eTiters are shown in dark blue and FCA titers are shown in green. (B) Comparison between eTiter and FCA titer. Titers of recombinant baculoviruses for the yeast Mediator Head module, yeast TFIIF, human Taf8-Taf10 heterodimer, yeast CycC-CDK8 complex, yeast core TFIIF, and the yeast Mediator Middle module were measured by two different methods and compared. Each measurement was repeated three times and averaged. eTiters are shown in dark blue and FCA titers are shown in green. (C) Different insect cell lines yield different eTiter values. The eTiters for a total of 12 viruses were measured in both Sf9 and Hi5 cells and compared. The eTiters in Sf9 are colored in dark blue, and those in Hi5 are colored in orange.

Figure 5. Quantification of expression of the multi-protein complexes involved in eukaryotic RNA polymerase II transcription by varying eI_{24} or eMOI values in the Hi5 and Sf9 cells. A total of six different multi-protein complexes were expressed in Hi5 cells and Sf9 cells under different conditions in terms of initial infectivity (eI_{24}) or eMOI ranging from $eI_{24} = 18\%$ (eMOI=0.2) to $eI_{24}=100\%$ (eMOI=4-5) in a 200 ml culture scale. SDS-PAGE of each purified complex with different eI_{24} or eMOI values is shown in right. Yield of complex at each point was measured and plotted against I_{24} (left) or eMOI (right) panel. The results from expression using Hi5 cells were shown on top and those from Sf9 cells were shown in the bottom. (A) Human Taf8-Taf10 complex, (B) yeast core TFIIF complex, (C) yeast TFIIF, (D) yeast CycC-CDK8, (E) the Mediator middle module, and (F) the Mediator Head module. Asterisk (*): contaminant

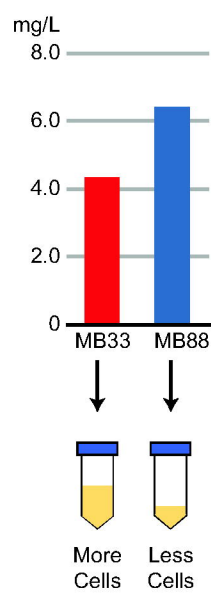
Figure 6. TEQC method ensures reproducibility of production of the multi-protein complexes. (A) Mediator Head module, Taf8-Taf10, TFIIF core, TFIIF, CycC-CDK8, and Middle were expressed at their optimum infectivity or eMOI, and each protein complex yield was measured; labeled as “initial” (dark blue). Expression was repeated three times using the same eMOI or infectivity value for each complex independently, purified protein complex yields were measured and the numbers were averaged; labeled as repeats. The repeats (green) were compared with the initial protein expression optimization results (dark blue). (B-E) Mediator Head, Taf8-Taf10, core TFIIF, and TFIIF complexes were expressed using freshly made new high titer viruses as well as the old viruses with declined titers. The eTiter values of both new and old viruses were determined and volume of each old or new virus was adjusted such that the expression of each complex was carried out at the same eI_{24} or eMOI. After expression and

purification, protein complex yield was measured, and the data were plotted against either eI₂₄ or eMOI. (B) Mediator Head module was expressed in Hi5 cells (red) (C) Taf8-Taf10 was expressed in Hi5 cells, (D) TFIIH core was expressed in Sf9 cells, (E) TFIIF was expressed in Sf9 cells, and compared with the initial expression results (blue) and the results from old viruses (red).

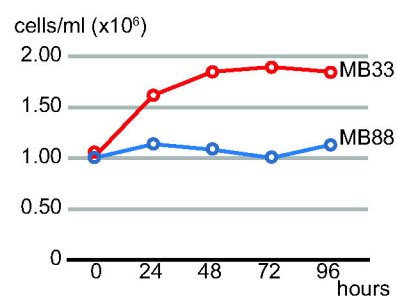
A

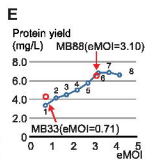
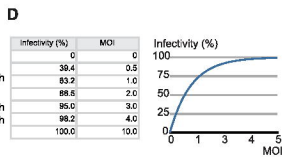
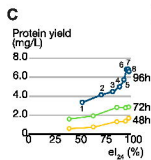
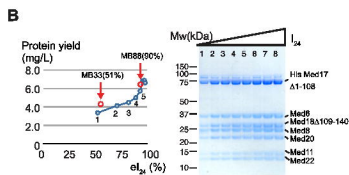
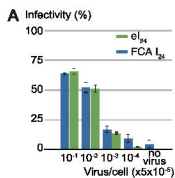


B



C

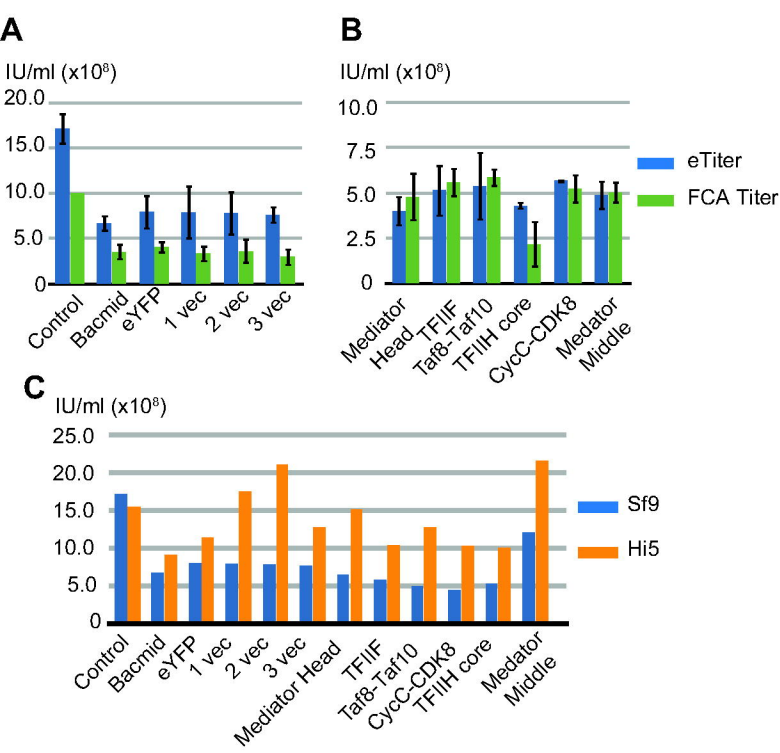




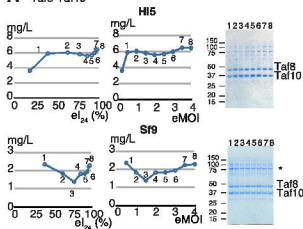
$$\begin{array}{c}
 \text{Protein} \\
 \text{complex} \\
 \text{production}
 \end{array}
 \longleftrightarrow
 \text{el}_{24} = 1 - e^{-\text{eMOI}} \quad (\text{Eq3})
 \longleftrightarrow
 \text{eTiter} = \frac{\text{eMOI} \times C_1 \times V_{\text{cell}}}{V_{\text{virus}}} \quad (\text{Eq5})$$

$$\uparrow$$

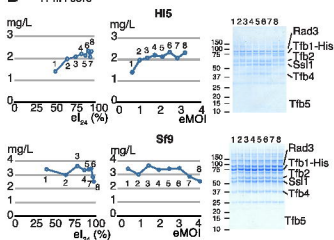
$$\text{el}_{24} = 1 - \frac{C_1 - C_0}{C_0(\alpha-1)} \quad (\text{Eq1})$$



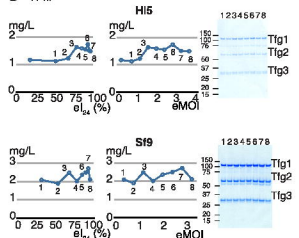
A Taf8-Taf10



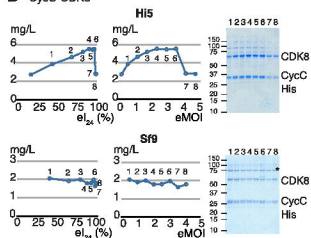
B TFIIH core



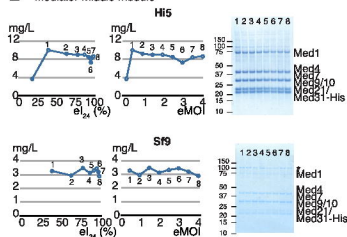
C TFIIIF



D CycC-CDK8



E Mediator Middle module



F Mediator Head module

