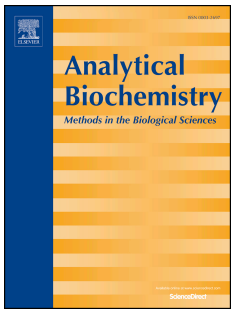
Accepted Manuscript



Long term kinetic measurements revealing precision and general performance of surface plasmon resonance biosensors

Franziska Steinicke, Imke Oltmann-Norden, Hermann Waetzig

PII: S0003-2697(17)30206-3

DOI: 10.1016/j.ab.2017.05.009

Reference: YABIO 12697

To appear in: *Analytical Biochemistry*

Received Date: 4 January 2017

Revised Date: 16 March 2017

Accepted Date: 8 May 2017

Please cite this article as: F. Steinicke, I. Oltmann-Norden, H. Waetzig, Long term kinetic measurements revealing precision and general performance of surface plasmon resonance biosensors, *Analytical Biochemistry* (2017), doi: 10.1016/j.ab.2017.05.009.

Title:

Long term kinetic measurements revealing precision and general performance of Surface Plasmon Resonance biosensors

Highlights:

1. Major error sources: surface phenomena and regeneration, and dilution steps
2. High precision can be achieved, e.g. approx. 3-6% RSD% for dissociation constants K_D
3. Control charts contribute to an overall better system performance

Abstract:

This work presents an extensive parameter list that facilitates a survey of biosensor performance using Biacore instruments for kinetic binding studies. Six long term measurements were performed using a strongly interacting antigen-antibody (β2 microglobulin) system. Both Single Cycle Kinetic (SCK) and Multi Cycle Kinetic (MCK) were executed each with five different analyte concentrations. The overall comparison of the long term monitored parameters, like the dissociation constant (K_D with approximately 3 to 6% relative percental standard deviation), the association and dissociation rate constants (k_a, k_d), the analyte binding capacity (R_{max}), chi^2 and the sum of the absolute values of the residuals, revealed the delicate factors that make the system performance vulnerable. The main influential factors on kinetic performance were the regeneration conditions, the quality of the sensor surface, the usage time and alteration of the sensor surface, the dilution series and the number of run cycles (about 250-600 per chip). Moreover the direct comparison of MCK and SCK uncovered distinct differences in the accuracy of the K_D values. The study of sensor chips from two manufacturers showed distinct differences in the precision of the data. Using control charts for the surveillance of these parameters contributes to an overall better system performance.

Keywords:

System performance; Surface Plasmon Resonance; Control charts; Human β2 microglobulin; Kinetic constants**;** Biosensors.

Abbreviations:

SCK, Single Cycle Kinetic; MCK, Multi Cycle Kinetic; K_D, dissociation constant; k_a, association rate constant; k_d, dissociation rate constant; R_{max}, surface analyte binding capacity; SPR, surface plasmon resonance; ICH, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; EU, European Union; GMP, good manufacturing practice; USP, United States Pharmacopoeia; HPLC, high performance liquid chromatography; CE, capillary electrophoresis; ITC, isothermal titration calorimetry; b2m, β2 microglobulin; RU, response units; IFC, integrated μ-fluidic cartridge; SD, standard deviation; RSD%, relative standard deviation in percent; SE, standard error; RI, bulk refractive index contributions in the samples; k_t, mass transfer constant; SOP, standard operational procedure.

1. Introduction:

1.1. Measuring drug-target interactions and Surface plasmon resonance analyses

A key step for the development of new drugs is the study of drug-target interactions by ligand binding assays. For characterization of biomolecular interaction and antigen-antibody interactions in particular, one of the most applied ligand binding assay employs surface plasmon resonance (SPR). SPR is a very suitable approach for measuring dissociation constants of drug-protein interactions because it covers a really broad range of dissociation constants that are accessible (K_D from $10^{-3.5}$ to $10^{-10.5}$ M). Moreover this technique facilitates label-free binding analysis studies of biomolecules such as affinity, kinetic, thermodynamics and specificity analysis in real-time [1,2]. In a typical SPR analysis setup one binding partner (the ligand) is bound to the surface and the other partner (the analyte) flows over the surface. The SPR sensor measures the refractive index change emerging through the association and dissociation of the analyte to the ligand.

Therefore SPR is an important application in pharmaceutical industry for drug discovery and proteomics especially from drug development to quality control and assurance [3]. Irrespective of the analytical technique qualified instruments and validated assays are an indispensible prerequisite for gaining precise and accurate results. Important guidelines like those of the ICH [4], EU GMP [5] and USP chapter 1058 [6] are the cornerstones in this field. Nevertheless there are, to the best of our knowledge, no publications dealing with analytical parameters of SPR instruments for performance qualification until now.

1.2. System performance

Existing validation reports and papers [7–10] have already established important parameters that have to be taken into account for biosensor assay development. The works of DiGiacomo [7] or Myszka [8] treat in detail issues in the development of new methods and assays and give important guidance to examine parameters like mass transport and its limitations, unspecific binding and baseline drifts. Based on these publications the present work outlines the key parameters for evaluating the performance of a SPR instrument. Leaving aside method development and validation it focuses on the most important parameters for performance qualification. The execution of long term kinetic measurements illustrates the different parameters which have a general influence on the system performance. Moreover these long term measurements decisively reveal covert tendencies or biases which are invisible in short term tests or normal use. Thus this study is working towards a general performance qualification method for surface plasmon resonance biosensors because this concept proved to be highly beneficial in the context of other analytical techniques such as HPLC [11] or CE [12]. For this purpose an already well established method was used in this study, the human β2 microglobulin antigen-antibody system [13,14]. Using an established method allows not to care about influences which hamper the performance but are mostly or exclusively associated with assay or method development. For example basic information like immobilization levels or suitable regeneration conditions are already checked out. Furthermore, the error sources matrix effects and aggregation can be prevented ab initio by using clean and purified sample material. Matrix effects and aggregation can certainly be major factors in SPR assays, but this important but complicated issue will need another comprehensive study.

Biacore devices have a system check option to control the basic functions of the instruments. Especially fundamental issues of the system like pump calibration and noise are monitored. This test can be performed by every Biacore user and it is recommended to retest it repeatedly. Furthermore in Biacore handbooks the fundamentals of the measurable parameters are well explained and give a good guidance in evaluating these parameters and

the quality of the measurements [13,15,16]. Moreover the evaluation software 2.0 of the Biacore X100 for example gives a first assessment of the measured parameters with a color coded ranking after each run. This tool gives a good and first evaluation of the used method and is a great support for method development and validation although it is not able to assess the performance of the instrument and especially not in long term measurements. As mentioned above no publications are concerning parameters that influence the qualification of performance in SPR instruments until now to our knowledge. An initial performance qualification option can only be executed by the technical support of the manufacturer (GE Healthcare) so far [17,18]. In this study for the first time long term SPR measurements were monitored by control charts which proved to be a powerful tool to reveal critical parameters of the system performance. Hence a monitoring system for the most important parameters such as k_a was developed using straightforward control charts.

1.3. The dissociation constant

The characterization of the drug-protein binding is usually performed by measuring the parameter dissociation constant K_D (or the reciprocal binding constant K_A respectively). K_D is directly dependent on the kinetic rate constants of a binding situation, more precisely it derives from the qoutient of the kinetic rate constants. In order to achieve maximum binding efficiency the K_D is often optimized for a class of substances in medicinal chemical investigations [19]. The K_D is regulated by the individual values of the rate constants, k_d and k_a. These rate constants can differ although their ratio (K_D) remains constant. Therefore the accurate measuring of the rate constants is a prerequisite for significant K_Ds. Furthermore K_D can be utilized to develop a deeper thermodynamically understanding in different binders by estimating the free energy or rather enthalpy and entropy [20]. For gaining the best results the used technique for measuring the dissociation constants plays an important role. So in the literature more than 20 very different techniques have been described for determining biomolecular binding kinetics of drug-protein interactions, such as equilibrium dialyses, ultrafiltration, ultracentrifugation, capillary electrophoresis techniques, spectroscopic techniques, calorimetric techniques and SPR [21,22]. Three of the most frequently used methods to measure the binding affinity of protein-target or especially protein-protein interactions are ITC, SPR and fluorescence based methods [23]. Past benchmark studies have explored the variability of kinetic rate constants across laboratories and instruments but (unfortunately) limited to SPR analytics [24–26].
The determination of binding affinity is clearly method depending since each method has its specific limitations [23]. Thus there is no single true K_D value and one cannot expect that different techniques provide identical dissociation constants [20]. Irrespective of the method an indispensible prerequisite for the determination of accurate and precise K_D values is a comprehensive system performance qualification. Therefore it is important to gain deep insight into the capability of the SPR system(s).

2. Materials and Methods:

2.1. Materials

Experiments were performed using a Biacore X100 from GE Healthcare (Uppsala, Sweden). Biosensor data produced were evaluated with Biacore evaluation software 2.0 in a 1:1 global fit. The mouse antibody against human β2 microglobulin (clone B2M-02, 150 kDa) and the β2 microglobulin (b2m, 11.8 kDa) from human urine were purchased from Sigma-Aldrich (Saint Louis, USA). Four CM5 sensor chips, the amine coupling kit containing 0.1 M *N*-

hydroxysuccinimide (NHS), 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and 1.0 M ethanolamine-HCl pH 8.5 and the maintenance products were obtained from GE Healthcare (Uppsala, Sweden). Two SPR Sensor chips CMD 500 l where purchased from Xantec bioanalytics. The regeneration solution 10 mM glycine-HCl pH 2.5 and the running buffer HBS-EP+ (HEPES 10 mM, NaCl 150 mM, EDTA 3 mM, polysorbate 20 0.05% v/v at pH 7.4) were prepared with ultrapure water and filtered with a nylon filter of 0.2 µm.

2.2. Immobilization, pre and post treatment

The immobilization and all following binding and kinetic measurements of anti-b2m antibody were performed at 25 °C and a flow rate of 30 µL/min. The protocol used was the immobilization and assay workflow of the Biacore Getting Started Kit [13] for kinetic measurements. The anti b2m antibody was immobilized to the sensor surface of each sensor chip using the amine coupling kit of GE Healthcare, resulting in a surface density of approximately 1200 response units (RU). One of the two flow cells was left blank for double-referencing the sensorgrams. After immobilization the maximum binding capacity was determined to monitor the activity of the bound antibody. Therefore three high concentrations of b2m (64, 128 and 256 nM in HBS-EP+) were measured to saturate the ligand binding sites on the surface and to determine the highest possible RU (real R_{max}). This test was repeated for each respective chip at the end of the set of kinetic measurements to monitor changes in real R_{max}.
Maintenance steps were performed to desorb and to sanitize the tubing and the flow cells (IFC) of the instrument as well as system check and pump calibration to monitor and maintain system status [8]. Desorbing was done at least every 30 runs, desorbing and sanitizing approximately every 60 runs directly followed by system check and pump calibration. When running maintenance the chips were stored in HBS-EP+ at 4 °C.

2.3. Single cycle kinetic (SCK) measurements

Five different analyte concentrations (2, 4, 8, 16 and 32 nM) were prepared with running buffer HBS-EP+ by subsequently diluting the stock solution of analyte (128 nM) as described in [13] for the SCK1 set. The serial dilution work flow was modified to a new pipetting schema for the SCK2 set though the concentrations remained the same. Hereupon every concentration was diluted individually from the stock solution. The workflow assays were implemented as recommended in [13] for kinetic measurements and for surface regeneration. Three SCK measurements were performed of one daily fresh serial dilution. In the SCK assay setup one start-up cycle to clear the surface is followed by two blind cycles for subtraction of baseline noise. Then the actual kinetic measurement follows in one cycle measuring the different analyte concentrations sequential in a titration way upwards to the highest. Every cycle is followed by one regeneration step of 30 seconds [27]. As a result can be stated that for one set of kinetic rate constants four cycles and four regenerations respectively have to be carried out.

2.4. Multi cycle kinetic (MCK) measurements

MCK assays start with two start-up cycles followed by one blind cycle. After this the analyte cycles (all concentrations in separate cycles) take place again followed by one blind cycle. In difference to SCK setup in MCK every cycle is associated with two regeneration steps for improved baseline return. The MCK1 set (/chip) was measured with a fixed sequence of analyte cycles (8, 4, 2, 16, 32 nM) whereas in the other set (/chip) of MCK2 five different sequences in a randomized pattern were implemented (*1.* 32, 16, 8, 4, 2 nM; *2.* 2, 4, 8, 16,

32 nM; *3.* 8, 4, 2, 16, 32 nM; *4.* 4, 32, 2, 16, 8 nM; *5.* 8, 4, 32, 2, 16 nM) [8]. According to SCK in MCK for one set of kinetic rate constants nine cycles have to be carried out at least and hence 18 regeneration steps because of two regeneration steps per cycle.

2.5. Evaluation

The parameters determined by the Biacore evaluation software 2.0 in kinetic evaluation on the basis of curve-fitting algorithms were extracted to Microsoft excel 2007 worksheets generated for these experiments. The arithmetic mean, the standard deviation (SD) and the relative standard deviation in percent (RSD %) were calculated for all values. The “raw” residuals were extracted from the residual plots of the evaluation software. The absolute values of these “raw” residuals were taken and summed up to create the sum of the absolute values of the residuals for every single kinetic measurement. All plots were generated with OriginPro 2016G.

3. Discussion and Results:

3.1. Parameters monitored

The purpose of this study was to identify the parameters that reflect best the instrumental performance of a Biacore X100 using a typical antigen-antibody binding study carried out by the SPR measurements. The parameters monitored for evaluating the performance were the association and dissociation rate constants (k_a and k_d), the dissociation constant (K_D) and the analyte binding capacity of the surface (R_{max}) all with their respective standard errors (SE). In addition the sum of the absolute values of the residuals from the fitted curve to the experimental data (sensorgram), the average squared residuals (chi^2), the bulk refractive index contributions in the samples (RI) and the mass transfer constant (k_t) were plotted.

All these parameters were analyzed in a long term kinetic study on one Biacore instrument to identify the most stable settings of a future PQ method. Secondly, it was analyzed which parameters indicate best the changes within the system, e.g. through the aging of the chip or rather of the bound antibody. In summary it is interesting to understand, which parameters react most sensitively to instrumental or external influences.

3.2. Analyte binding capacity (R_{max})/ Surface performance

The analyte binding capacity of the ligand (R_{max}) is a global parameter directly given by the Biacore evaluation software. It is useful to continuously monitor the binding capacity over the time period the chips are in use, in order to get information about the condition and lifetime of one chip and thus about changes of surface performance. The R_{max} represents directly changes of the efficiency of the analyte binding capacity over the whole long term study due to the covalent immobilization of the antibody on the surface.
The regeneration step is one of the essential influences on the analyte binding capacity. An example for the critical effect of inefficient chip regeneration is illustrated by the comparison of the long-term MCK measurements of MCK1 and MCK2 (cf. Fig. 2). For each test series one chip was used, respectively. In MCK1 the values of R_{max} remain steady over all measurements and only decrease throughout the usage time with increasing number of runs. However in MCK2 the R_{max} shows much stronger variations over the whole number of runs generated by the varying sequences of the sample concentrations in each measurement (cf. section 2.4.). In measurements starting with higher concentrated samples the regeneration steps were mostly not sufficient and the R_{max} value increased, unlike to measurements

starting with lower concentrations (cf. Fig. 1). Interestingly the all over decrease of R_{max} was similar to the one of MCK1 over the whole usage time of the MCK2 chip. Harsher regeneration conditions can have an unwanted effect on the protein structure of the covalently bound ligand contrary to slightly insufficient regeneration conditions. Harsher conditions can destroy the activity of the ligand very quickly. Hence it was necessary to improve the inefficient regeneration condition but without affecting the activity of the ligand too much. A suitable solution to this challenge was to add a repetitive regeneration step and not to change the regeneration buffer. Changing to harsher conditions like lower or much higher pH leads to too many differences between the setups. Variances in the experimental setups should be avoided in order to preserve the comparability between the experiments. Figure 1 shows the importance of the second regeneration step. In the MCK setup every concentration is measured in a separate cycle. Therefore no start-up cycles are taking place, which would have an additional cleaning effect on the surface, before the second to fifth concentrations are measured. Remaining analyte on the surface exhibited major effects on the obtained data if the surface was not cleaned with a second repetetive regeneration step of Glycine 10 mM pH 2.5 for 30 seconds.Thus this example of sensor surafce regeneration problems evidently reveals that the R_{max} reflects very good the efficiency of the regeneration conditions and steps.

In comparison to SCK in MCK measurements the analyte binding capacity R_{max} did not decrease proportionally to the number of regeneration steps. Although the MCK setup afforded five times more regeneration steps per set of kinetic measurements the decrease of R_{max} was not more pronounced than under MCK conditions. This indicates that the total number of regeneration steps has a minor influence on the analyte binding capacity. The time period the chips are used including storage time and run cycles is probably the major factor for influencing the surface performance (Fig. 2).

In this study it can be observed in Figure 2 that SCKs performed on the two chips from Xantec offer a higher immobilization level although using the same immobilization assay in these two measurement setups. Furthermore they show particularly constant values without any major variations. Hence the composition of the chip surface turned out to be a further influential factor on R_{max}. Both kinds of chips are declared to be produced with the same surface properties in length of the dextran chains and occurrence of the carboxymethyl groups by the manufacturers. Thus relating to manufacturers' data the Biacore and the Xantec chips are supposed to behave equally. However, they are different not only in their constancy of values but also in their immobilization levels. The manufacturing method is obviously not the same and this can have an effect on the performance. Accordingly the chip properties have a great impact on the method validation. However to measure the performance of the instrument for qualification one is dependent on the R_{max} and therefore the sort of chips has an impact on the PQ too.

The four Biacore chips were immobilized with R_{max} levels between 110 and 80 RU whereas the Xantec chips show a higher and possibly more stable level of approximately 120 RU. Nevertheless a disadvantage of the Xantec chips in contrast to the original Biacore chips is the less user friendly storage behavior. Biacore chips can easily be stored in HBS-EP+ buffer or other running buffers at 4 °C. If anything the chips lose a little bit of active antibody from the surface (cp. Figure 2 where the chip of SCK2 was stored in buffer in the fridge after run 35 for 4 weeks and the R_{max} scarcely decreases). Whereas the Xantec chips have to be stored dry and under inert gas atmosphere (with a droplet of stabilization buffer on the gold surface) because of the hydrogel coated surface which possibly could be hydrated when stored in buffer [28]. This hydrated hydrogel surface can then potentially stick to the optical interface of the instrument and damage it when the chip is docked. In Figure 2 (and Figure 5)

the poor storage behavior is represented in the SCK Xantec 1 data after run 50 where the chip was stored dry under argon atmosphere in the fridge for two days. The decrease of the R_{max} during this interruption was about 14 RU. This decrease was not influencial for the kinetic values probably because of the initially higher immobilization level but it is important to monitor such alterations to estimate the long time performance and life time of the chip and the significance of all the values obtained. To reduce this unfavorable storage effect the manufacturer recommends a special stabilization buffer to be dropped on the surface of the chip while storing under dry conditions to keep the bound ligand hydrated and active [28]. Additionally to the storage behaviour Figure 2 exemplifies the scatter in the R_{max} value, in particular with regards to the SCK2 and MCK2 where the values spread broader in general. However Figures 6 and 7 prove that R_{max} is of no substantial influence on the kinetic data. Even after 70 runs with R_{max} down to half of its initial value only a moderate increase of k_a and k_d can be seen. Nevertheless it is important to monitor the R_{max} together with the kinetic data because changes on the surface of the chip can then easily be recognized. Changes of the ligand on the chip surface have an influence on the kinetic rate constants because these become more variable at the final stages of the chip-lifetime. In fact, R_{max} is the only variable which reflects this. Thus the R_{max} value is a definitely suitable parameter for the surveillance of the chip surface over the whole chip-lifetime.

3.3. Dissociation rate constant (k_d)

The kinetic values k_d and k_a and the resulting K_D from the ratio of k_d/k_a are extracted from the experimental curve by the Biacore evaluation software with a fitting algorithm. Observing the experiments over the course of time it became apparent that k_d was the more stable and independent parameter to monitor in this work (cf. Fig. 5). Over all long term measurements the RSD % of the k_d was between 2.5 and 6.2 % over a period of at least 30 up to 90 kinetic measurements for each tested chip.
The dissociation rate constant is almost completely independent of external influences. As a consequence k_d is a very useful value to measure the system suitability. On the contrary k_d is unsuitable to be monitored in long term measurements to obtain information about the external influences.
In the plot of k_d of SCK Xantec 1 it is possible to observe a slight but distinctive increase of the value over time (Figure 6). This increase could indicate a slow alteration of the surface of the chip by e.g. the bound antibody. Probably the regeneration steps detach some antibodies and the remaining antibodies are easier to be assessed and departed such that the k_d increases as well as the association rate constant k_a (cf. Figure 6). However k_a is further influenced by several parameters such as the dilution protocol (section 3.4.). This will be discussed further in section 3.10.
For kinetic measurements it is recommended to immobilize as little as possible [8] but as much as necessary such that the signal does not disappear in the baseline noise. In this study the reduction of the immobilization level was discussed as a further approach. The reduction of the immobilization level is common practice for optimizing kinetic measurements during early method development. It was rejected here because our aim was to perform this study with a well-established assay and because the employed immobilization level proved to be very well acceptable. Any changes would have decreased the comparability between the past and future measurements using this prominent assay.

3.4. Association rate constant (k_a)

In contrast to k_d association rate constant k_a was most suitable for monitoring the instrumental and external influences because it turned out to be most sensitive for any variations of the measurement process. This became apparent in the control charts of k_a of all chips (cf. Figures 3 and 4 as example). For example the exchange of vials for the dilution series from one manufacturer to another increased the spreading of the values in k_a of SCK1 (cf. Figure 3). In this context the influence of leaching from plastic disposables has to be taken into account [29,30]. The substances leaching from plastic vials in the aqueous buffer can affect the protein function or surface/interface tension of the solution and therefore the outcome of the kinetic assay. Different producer of plastic vials can introduce diverse leachables in different amounts and therefore influence the kinetic values in various ways. For this reason the vials were changed back to the first vial type. Further modifications of k_a caused by external influences can be seen in Figure 4 in the control chart of SCK Xantec 1, but also in Figures 3 and 6. Firstly the Figures show an incontrovertibly strong correlation between the k_a and the dilution series which appears in the grouping of three connected k_a data points each. This correlation arises from the measurement of always three respective runs out of one dilution series each day. If a mistake or a slight variation occurs in the preparation of the dilution series [31–33] the parameter k_a will reveal this. This effect is observable in R_{max} (cf. Figure 2), too, but milder in its appearance. Secondly k_a in SCK Xantec 1 (Figure 4) illustrates well the drift upwards of k_a over the time as aforementioned (in section 3.3.). Thirdly the more strongly varying k_a values in SCK Xantec 1 after run 50 indicate that the handling operator is a very powerful influence even though he is widely experienced and is following a standard operating procedure (SOP). The RSD%s of k_a spread enormously over all long term measurements between 4 and 35 %. Presumably a person can not reach the accurateness of for example pipetting robots but with longer experience one can gain a very high precision. Established procedures as well can never completely avoid deviations caused by unwritten aspects like exact movements of the hand for every working step which pose a potential for spreading values. Nevertheless SOPs are a very helpful tool and indispensable in every lab because of their harmonization of the work flow.

3.5. Dissociation constant (K_D)

K_D shows a comparable fluctuation range as k_a but is normalized through k_d which is more stable (cf. Figure 8 and section 1). At this point it is important to distinguish between precision and accuracy. Precision is defined as the measure of random errors respectively the spreading of repetitive measurements of matching experimental setups, whereas the accuracy describes the fluctuation around the true value of repetitive measurements of matching experimental setups. This can be explained on the basis of the MCK measurements. The only difference between MCK1 and MCK2 was the varying sequence of the differently concentrated analyte samples. In MCK1 the value of K_D has a RSD % of 8.31 % which exhibits just some spreading, or in other words the values are exceptionally close to each other. Accordingly the precision of the K_D is quite good. This precision can be related to the fixed sample injection sequence. The fixed sequence smoothens the deviation of the occuring values because every time it is the same last concentration in the fixed sequence which influences the most the unefficient regeneration (cp. section 3.2). This results in an error of always the same order/magnitude. Whereas in MCK2 (the series with random sample sequence) the values of K_D spread more widely and therefore have a RSD % of 15.1 % revealing a worse precision. The reason for this is that in MCK2 every sample sequence exhibits its own error as a result of the differing residue of analyte after every run.

Therefore the K_D values deviate more. However MCK2 describes the measurement error more realistically because the variability due to the regeneration step is considered better for the β2 microglobulin antigen-antibody system (Figure 7). This is why MCK2 is to be supposed less precise but more accurate in the average than MCK1. Figure 7 shows that MCK1 is more precise and possibly less accurate in comparison to MCK2 because its K_Ds are on the bottom line of all of our measured K_Ds. Accuracy is a lot more difficult to assess compared to precision. In order to determine the exact true value interlaboratory trials [24] with diverse techniques [21,25] and operators are required, as mentioned in the introduction.

3.6. chi^2

Chi squared (χ^2; chi^2) is a parameter depicting the goodness of fit. It is used in the Biacore evaluation software to describe the differences from the experimental data of the sensorgram to the fitted curve.

$$chi^2 = \frac{\sum_1^n (r_f - r_x)^2}{n - p}$$

Equation 1: chi^2

The threshold up to which chi^2 is accepted needs to be defined before starting the measurements. The threshold limit value depends on the analyte ligand system used but a value ≤ 10 is acceptable in most of the cases [34]. In the present work the worst chi^2 value was about 3.07 (in SCK 1 in run 14) in our set of experimental data. The chi^2 was between 0.2 and 0.8 in the average of all sets of kinetic measurements. Thus in this system the chi^2 is very far below the mostly requested threshold. Nevertheless the MCK setup provided lower χ^2 values (MCK1 0.2, MCK2 0.4) than the SCK setup (SCK1 0.5, SCK2 0.8, SCK Xantec 1 and 2 0.6). The chi^2 depicts very well changes of single measurements in numerical values for this antigen-antibody system. Thus the increase of chi^2 above the average leads also to an evident change of the kinetic data. For example the comparison of the runs 13, 31, 79 and 84 in Figures 3 and 8 show the changes of the kinetic data (k_a) and the parallel rise of the chi^2 value which demonstrate the strong correlation of both parameters. Therefore it is a valuable indicator to depict changes in this study. Admittedly the original shape of the curve is in some cases more important than the chi^2 value itself. Hence it is always recommendable to (re)consider the sensorgram i.e. the curve progression and not to underestimate its importance.

3.7. Sum of the absolute values of the residuals

The sum of the absolute values of the residuals was monitored for each kinetic measurement and for each individual chip in the long-term measurement series. Moreover the sum of the absolute values of the residuals can be calculated for every single concentration that was measured. Thus a plot can be generated of the sum of the absolute values of the residuals for each concentration in a long-term study. The possibility to monitor single concentrations as well as the sum of the absolute values of the residuals enabled to observe small deviations in the dilution series. Furthermore it revealed which concentration showed the largest deviations over all measurements or in every single run over the whole time course. In this study the lowest values for the sum of the absolute values of residuals for SCK and MCK were found at the concentrations of 8 nM and 32 nM, although the value for 32 nM included the 600 seconds of dissociation time in SCK measurements. It was striking that even in MCK2 with a random sequence of measured concentrations those two showed the lowest discrepancy. This value is a supplemental parameter to illustrate the goodness of fit

between the experimental curve and the fitted curve. Therefore it shows the comparable discrepancies as the chi^2 but is calculated differently (cf. section 2.). In this study it was helpful to have a further parameter to monitor the long term developing of the values. Furthermore this parameter was used to validate the chi^2 in our measurements, but for further works it is possibly nonessential.

3.8. Mass transfer constant (k_t)

In case of the β2 microglobulin test system and the MCK measurements k_t was about 10^8 $RU*M^{-1}*s^{-1}$ and hence in the accepted range [15]. Whereas in the SCK measurements k_t yielded 10^{21} $RU*M^{-1}*s^{-1}$.This value is so unreasonably high that it indicates no mass transfer limitations at all for this measurement system. Often the k_t is a parameter with rather low significance because it can cover a wide range of values without effecting the closeness of the fit [15]. Nevertheless it has an important control function because if measurements do not work as expected and the k_t values do not keep the limit this will be a crucial indication that the system is mass transfer limited [15,34]. Thus the mass transfer effect should be controlled in the step of method development.

3.9. Bulk refractive index contribution (RI)

The RI is directly dependent on the optical refractive index of the analyte solution. In turn the optical refractive index depends on the concentration of the analyte and the buffer constitution. The RI was a quite stable parameter in this study though it indicated outliners as well as the chi^2 in some cases in parallel. This indicates that some outliners which are visible in R_{max}, the kinetic values, chi^2 and the RI are caused by dilution errors. In contrary to those outliners that are only visible in R_{max}, the kinetic values and chi^2. These were more probably induced by other external or internal influences as the dilution series.

3.10. MCK and SCK in comparison

The first set of MCK measurements was performed with a fixed order of analyte concentration whereas the second set was performed in a randomized pattern. The intention of the first set was to create a setup as comparable as possible to the SCK measurements to minimize the discrepancy of the data collected. In contrast, however, the values of the kinetic constants differed more between the SCKs and the MCK1 with fixed order than between the SCKs and the MCK2 with random injection sequence.
Furthermore the phenomenon of increasing kinetic rate constant became visible over the long term measurements. Especially in MCK2 when comparing the five different sequences individually for the first time a slight but incontrovertible increase in the k_d for almost every sequence becomes visible (Figure 9). This implements that the value independently increases from the concentration series and presumably is dependent on the life time of the chip and on the density and age of the bound antibody, respectively.

4. Conclusions:

A comprehensive list of performance parameters for kinetic analysis on a Biacore X100 instrument has been provided using a well established method. The antigen-antibody system used was confirmed to be very useful for long term measurements because of its robust storage and handling behavior and good availability. In particular its kinetic values are easily accessible and well within the typical measurement range.
As main parameters R_{max}, k_a and k_d as well as the chi^2 (or alternatively the sum of the absolute values of the residuals) have been identified. Furthermore the influence of RI and k_t

has been discussed. For optimal performance the sensor surface, the dilution series and the training of personnel deserve particular attention. Surface regeneration plays a crucial role. Surface materials of two manufacturers have been investigated in this study, and considerable differences have been found. Significant aging effects have been observed for both chip materials. The R_{max} is particularly suited to trace these. Interestingly the life time of a chip is more dependent on the time it is actually used and stored than on the number of regeneration cycles. However, the R_{max} can go down to half of its initial value without having negative effects on the kinetic rate constants. There are several pitfalls organizing dilution series; a good practice is the individual one step dilution of each concentration as described in the experimental part (section 2.3.). Standard operational procedures and periodic maintenance of the instrument are important for personnel training. When comparing SCK and MCK measurements, both showed up with comparable advantages and disadvantages for a future performance qualification method. For example SCK was a more time saving experimental setup (with SCK one result was obtained in 3 hours, with MCK in 4 hours) but the MCK setup did not exhibit as many outliners as SCK.

Visually displaying or plotting data is a prerequisite for correct interpretation [35] of every set of data but in particular for performance qualification. This work pinpoints that most effects firstly become obvious with visually comparing all data against each other. Every handling error or slight variation becomes manifest in the control charts [36–38].

This work provides reference data to compare with own measurements. Furthermore, this concept of the evaluation on SPR instruments can be generalized. This is an important first step towards a thorough SPR performance qualification. A corresponding concept will be outlined in a follow-up work [39]. An RSD% of 3-6% for K_D can be achieved for long measurement series ($n \geq 60$).

Acknowledgements:

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

We thank Dr. Luma Baydoun and the Fraunhofer ITEM for providing the Biacore X100 instrument used in this study.

We are grateful to GE Healthcare Life Sciences for excellent technical support.

We thank Dr. Lutz Preu for proof reading this article.

References

[1] R.L. Rich, D.G. Myszka, Grading the commercial optical biosensor literature-Class of 2008: 'The Mighty Binders', J. Mol. Recognit. 23 (2010) 1–64.

[2] R.L. Rich, D.G. Myszka, Survey of the 2009 commercial optical biosensor literature, J. Mol. Recognit. 24 (2011) 892–914.

[3] J.-P. Renaud, C.-W. Chung, U.H. Danielson, U. Egner, M. Hennig, R.E. Hubbard, H. Nar, Biophysics in drug discovery: impact, challenges and opportunities, Nat. Rev. Drug Discov. 15 (2016) 679–698.

[4] ICH Harmonised Tripartite Guideline, Validation of Analytical Procedures: Text and Methodology Q2(R1), 2005.

[5] European Commision, EU Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterenary Use. EudraLex Volume 4 Annex 15, 2015.

[6] USP 37, Analytical Instrument Qualification General Chapter<1058>.

[7] R.A. DiGiacomo, L. Xie, C. Cullen, S.R. Indelicato, Development and validation of a kinetic assay for analysis of anti-human interleukin-5 monoclonal antibody (SCH 55700) and human interleukin-5 interactions using surface plasmon resonance, Anal. Biochem. 327 (2004) 165–175.

[8] D.G. Myszka, Improving biosensor analysis, J. Mol. Recognit. (1999) 279–284.

[9] R.L. Wong, D. Mytych, S. Jacobs, R. Bordens, S.J. Swanson, Validation parameters for a novel biosensor assay which simultaneously measures serum concentrations of a humanized monoclonal antibody and detects induced antibodies, J. Immunol. Methods (1997) 1–15.

[10] B. DeSilva, W. Smith, R. Weiner, M. Kelley, J. Smolec, B. Lee, M. Khan, R. Tacey, H. Hill, A. Celniker, Recommendations for the Bioanalytical Method Validation of Ligand-Binding Assays to Support Pharmacokinetic Assessments of Macromolecules, Pharm. Res. 20 (2003) 1885–1900.

[11] L. Kaminski, M. Degenhardt, J. Ermer, C. Feussner, H. Höwer-Fritzen, P. Link, B. Renger, M. Tegtmeier, H. Wätzig, Efficient and economic HPLC performance qualification, J. Pharm. Biomed. Anal. 51 (2010) 557–564.

[12] C. Cianciulli, H. Wätzig, Analytical instrument qualification in capillary electrophoresis, Electrophoresis 33 (2012) 1499–1508.

[13] GE Healthcare Lifesciences (Biacore), Biacore X100 Getting Started, GE Company, 2009.

[14] H. Zhao, I.I. Gorshkova, G.L. Fu, P. Schuck, A comparison of binding surfaces for SPR biosensing using an antibody-antigen system and affinity distribution analysis, Methods 59 (2013) 328–335.

[15] GE Healthcare Lifesciences (Biacore), Biacore X100 Handbook, GE Company, 2008.

[16] GE Healthcare Lifesciences (Biacore), Sensor Surface Handbook. BR-1005-71, GE Company, Uppsala, 2008.

[17] GE Healthcare Lifesciences (Biacore), Biacore T200 GxP Package, 2010.

[18] GE Healthcare Lifesciences (Biacore), Biacore C, 2013.

[19] H. Wätzig, I. Oltmann-Norden, F. Steinicke, H.A. Alhazmi, M. Nachbar, D.A. El-Hady, H.M. Albishri, K. Baumann, T. Exner, F.M. Böckler, S. El Deeb, Data quality in drug discovery: the role of analytical performance in ligand binding assays, J. Comput. Aided Mol. Des. 29 (2015) 847–865.

[20] C. Kramer, T. Kalliokoski, P. Gedeck, A. Vulpetti, The experimental uncertainty of heterogeneous public K(i) data, J. Med. Chem. 55 (2012) 5165–5173.

[21] K. Vuignier, J. Schappler, J.-L. Veuthey, P.-A. Carrupt, S. Martel, Drug-protein binding: a critical review of analytical tools, Anal. Bioanal. Chem. 398 (2010) 53–66.

[22] Mia Yang (HN), Biosensors in the Processing and Analysis of Biopharmaceuticals, in: G. Subramanian (Ed.) Biopharmaceutical Production Technology, Wiley-VCH, s.l., 2012 pp. 445–471.

[23] P.L. Kastritis, Bonvin, A. M. J. J., On the binding affinity of macromolecular interactions: daring to ask why proteins interact, J. R. Soc. Interface 10 (2012) 20120835.

[24] R.L. Rich, G.A. Papalia, P.J. Flynn, J. Furneisen, J. Quinn, J.S. Klein, P.S. Katsamba, M.B. Waddell, M. Scott, J. Thompson, J. Berlier, S. Corry, M. Baltzinger, G. Zeder-Lutz, A. Schoenemann, A. Clabbers, S. Wieckowski, M.M. Murphy, P. Page, T.E. Ryan, J. Duffner, T. Ganguly, J. Corbin, S. Gautam, G. Anderluh, A. Bavdek, D. Reichmann, S.P. Yadav, E. Hommema, E. Pol, A. Drake, S. Klakamp, T. Chapman, D. Kernaghan, K. Miller, J. Schuman, K. Lindquist, K. Herlihy, M.B. Murphy, R. Bohnsack, B. Andrien, P. Brandani, D. Terwey, R. Millican, R.J. Darling, L. Wang, Q. Carter, J. Dotzlaf, J. Lopez-Sagaseta, I. Campbell, P. Torreri, S. Hoos, P. England, Y. Liu, Y. Abdiche, D. Malashock, A. Pinkerton, M. Wong, E. Lafer, C. Hinck, K. Thompson, C. Di Primo, A. Joyce, J. Brooks, F. Torta, Bagge Hagel, Anne Birgitte, J. Krarup, J. Pass, M. Ferreira, S. Shikov, M. Mikolajczyk, Y. Abe, G. Barbato, A.M. Giannetti, G. Krishnamoorthy, B. Beusink, D. Satpaev, T. Tsang, E. Fang, J. Partridge, S. Brohawn, J. Horn, O. Pritsch, G. Obal, S. Nilapwar, B. Busby, G. Gutierrez-Sanchez, R.D. Gupta, S. Canepa, K. Witte, Z. Nikolovska-Coleska, Y.H. Cho, R. D'Agata, K. Schlick, R. Calvert, E.M. Munoz, M.J. Hernaiz, T. Bravman, M. Dines, M.-H. Yang, A. Puskas, E. Boni, J. Li, M. Wear, A. Grinberg, J. Baardsnes, O. Dolezal, M. Gainey, H. Anderson, J. Peng, M. Lewis, P. Spies, Q. Trinh, S. Bibikov, J. Raymond, M. Yousef, V. Chandrasekaran, Y. Feng, A. Emerick, S. Mundodo, R. Guimaraes, K. McGirr, Y.-J. Li, H. Hughes, H. Mantz, R. Skrabana, M. Witmer, J. Ballard, L. Martin, P. Skladal, G. Korza, I. Laird-Offringa, C.S. Lee, A. Khadir, F. Podlaski, P. Neuner, J. Rothacker, A. Rafique, N. Dankbar, P. Kainz, E. Gedig, M. Vuyisich, C. Boozer, N. Ly, M. Toews, A. Uren, O. Kalyuzhniy, K. Lewis, E. Chomey, B.J. Pak, D.G. Myszka, A global benchmark study using affinity-based biosensors, Anal. Biochem. 386 (2009) 194–216.

[25] D. Yang, A. Singh, H. Wu, R. Kroe-Barrett, Comparison of biosensor platforms in the evaluation of high affinity antibody-antigen binding kinetics, Anal. Biochem. 508 (2016) 78–96.

[26] P.S. Katsamba, I. Navratilova, M. Calderon-Cacia, L. Fan, K. Thornton, M. Zhu, T.V. Bos, C. Forte, Della Friend, I. Laird-Offringa, G. Tavares, J. Whatley, E. Shi, A. Widom, K.C. Lindquist, S. Klakamp, A. Drake, D. Bohmann, M. Roell, L. Rose, J. Dorocke, B. Roth, B. Luginbühl, D.G. Myszka, Kinetic analysis of a high-affinity antibody/antigen interaction performed by multiple Biacore users, Anal. Biochem. 352 (2006) 208–221.

[27] R. Karlsson, P.S. Katsamba, H. Nordin, E. Pol, D.G. Myszka, Analyzing a kinetic titration series using affinity biosensors, Anal. Biochem. 349 (2006) 136–147.

[28] Xantec bioanalytics, Product information: Immobilization on carboxylated and NHS activated surfaces, 2015.

[29] A. Olivieri, O.S. Degenhardt, G.R. McDonald, D. Narang, I.M. Paulsen, J.L. Kozuska, A. Holt, On the disruption of biochemical and biological assays by chemicals leaching from disposable laboratory plasticware, Can. J. Physiol. Pharmacol. 90 (2012) 697–703.

[30] G.R. McDonald, A.L. Hudson, S.M.J. Dunn, H. You, G.B. Baker, R.M. Whittal, J.W. Martin, A. Jha, D.E. Edmondson, A. Holt, Bioactive Contaminants Leach from Disposable Laboratory Plasticware, Science 322 (2008) 917.

[31] A. Petzold, A. Altintas, L. Andreoni, A. Bartos, A. Berthele, M.A. Blankenstein, L. Buee, M. Castellazzi, S. Cepok, M. Comabella, C.S. Constantinescu, F. Deisenhammer, G. Deniz, G. Erten, M. Espino, E. Fainardi, D. Franciotta, M.S. Freedman, V. Giedraitis, N.E. Gilhus, G. Giovannoni, A. Glabinski, P. Grieb, H.-P. Hartung, B. Hemmer, S.-K. Herukka, R. Hintzen, M. Ingelsson, S. Jackson, S. Jacobsen, N. Jafari, M. Jalosinski, S. Jarius, E. Kapaki, B.C. Kieseier, M.J.A. Koel-Simmelink, J. Kornhuber, J. Kuhle, J. Kurzepa, P.H. Lalive, L. Lannfelt, V. Lehmensiek, P. Lewczuk, P. Livrea, F.

571 Marnetto, D. Martino, T. Menge, N. Norgren, E. Papuc, G.P. Paraskevas, T. Pirttila, C. Rajda, K.
572 Rejdak, J. Ricny, D. Ripova, L. Rosengren, M. Ruggieri, S. Schraen, G. Shaw, C. Sindic, A. Siva, T.
573 Stigbrand, I. Stonebridge, B. Topcular, M. Trojano, H. Tumani, H.A.M. Twaalfhoven, L. Vecsei, V.
574 van Pesch, H. Vanderstichele, C. Vedeler, M.M. Verbeek, L.M. Villar, R. Weissert, B. Wildemann,
575 C. Yang, K. Yao, C.E. Teunissen, Neurofilament ELISA validation, J. Immunol. Methods 352 (2010)
576 23–31.

577 [32] N.G. Hentz, T.R. Knaide, Effect of liquid-handling accuracy on assay performance, J. Lab.
578 Automation 19 (2014) 153–162.

579 [33] K. Pandya, C.A. Ray, L. Brunner, J. Wang, J.W. Lee, B. DeSilva, Strategies to minimize
580 variability and bias associated with manual pipetting in ligand binding assays to assure data
581 quality of protein therapeutic quantification, J. Pharm. Biomed. Anal. 53 (2010) 623–630.

582 [34] GE Healthcare Lifesciences (Biacore), BIAevaluation Software Handbook, 2008.

583 [35] A. Cornish-Bowden, Detection of errors of interpretation in experiments in enzyme kinetics,
584 Methods 24 (2001) 181–190.

585 [36] J.C. Benneyan, Use and interpretation of statistical quality control charts, Int. J. Qual. Health
586 Care 10 (1998) 69–73.

587 [37] E. Mullins, Introduction to Control Charts in the Analytical Laboratory - Tutorial Review,
588 Analyst (1994) 369–375.

589 [38] G. Wernimont, Use of Control Charts in the Analytical Laboratory, Ind. Engin. Chem. 18
590 (1946) 587–592.

591 [39] F. Steinicke, I. Oltmann-Norden, H. Wätzig, Performance Qualification of a Surface Plasmon
592 Resonance Biosensor, Anal. Biochem. (2017) in preparation.

593

Table of Figures:

Figure 1: Left: Example for a SCK measurement sensorgram depicting the poor baseline return after only one regeneration. This incomplete return was improved by a start-up cycle in the next measurement. This start-up cycle cleans the surface almost completely and is not part of the analytical evaluation in the kinetic measurement. Measurements were performed as described in “SCK methods“. Right: Typical MCK measurement sensorgram with poor baseline return after the first regeneration step. Measurements performed as described in “MCK measurements”.

Figure 2: Decrease of the analyte binding capacity (R_{max}) over time. **SCK1**: Three measurements a day, one regeneration step after every cycle, only short term storage in buffer for maintenance procedures. **SCK2**: Three measurements a day, three times stored in buffer at 4 °C (after run 9 for 3 days, after run 35 for 4 weeks, after run 64 for 3 ½ weeks). **MCK1**: Two measurements a day, two regeneration steps after every cycle, twice stored in buffer at 4 °C (for 3 days each after run 32 and 40). **MCK2**: Three measurements a day, two regeneration steps after every cycle, five varying sequences of the injected concentrations, three times stored in buffer at 4 °C (for 3 weeks after run 6, for 10 days each after run 29 and 42). **SCK Xantec 1**: Three measurements a day, one regeneration step, once stored in buffer at 4 °C (for 2 days after run 50). **SCK Xantec 2**: Three measurements a day, two regeneration steps.

Figure 3: Control chart of the association rate constant of the SCK1 (Biacore sensor chip) measurement, relating events to the measured values. Green arrow: exchange of buffer

Figure 4: Control chart of the association rate constant of the SCK Xantec 1 measurement with a chip from the manufacturer Xantec.

Figure 5: A: Association rate constants of all long term measurements; B: Dissociation rate constants of all long term measurements; (Legend: cf. Figure 2)

Figure 6: SCK Xantec 1 A: Association rate constant with standard error (SE); B: Dissociation rate constant with SE; For closer examination of the small SE both are shown with magnified insert with comparable scale of three data points each.

Figure 7: A: Overview of all dissociation constants for each kinetic test series. The mean values as well as the standard deviations differ from series to series. B: Dissociation constant represented by the K_D plot. Dissociation rate constants are plotted against their corresponding association rate constants on logarithmic scale. For each chip the first 60 runs are displayed except for SCK Xantec 2 with only 31 runs. Two reference lines for the KD values $1x10^{-9}$ and $2x10^{-9}$ M are given also to show the coherence of all measured values.

Figure 8: Illustration of the strong correlation of chi^2 and the sum of the absolute values of residuals on the basis of SCK 1. chi^2 as some sort of variance and the sum of the absolute values of the residuals are directly comparable.

Figure 9: Dissociation rate constant depicted over number of runs in MCK2, ordered by sequence (not by actual run order). The main effect is due to the usage time (aging of the sensor surface) and not to the chosen sequence.