# Blivion

Maria J Schilstra & Stephen R martin

Structural Biology Science Technology Platform

Francis Crick Institute

1 Midland Road, London N1 1AT, UK

## Purpose

Blivion is a software tool for analysing biolayer interferometry measurements of the binding of large vesicles such as influenza virus onto a two-dimensional ligand-coated substrate. A typical experiment is carried out with eight probes and has the following phases:

1. *Baseline*: establishment of the baseline response
2. *Loading*: loading of the ligand onto the probe surface
3. *Loaded*:establishment of the ligand-loaded response
4. *Association*: binding of the vesicles to the coated surface
5. *Dissociation*: dissociation of the vesicles from the surface

In Phase 1, the response of the uncoated surface of the probes is measured. In phase 2, the substrate is coated to different extents by incubating the probes for an appropriate period with a range of ligand concentrations. At the start of phase 3, the ligand solution is removed, and the newly coated surface is allowed to stabilize. After the coated substrate has stabilized, phase 4 is started by dipping it into a vesicle solution; the same vesicle concentration for each probe. Vesicle binding occurs on a timescale of tens or hundreds of seconds, and is allowed to progress some way toward completion. Phase 5, in which the vesicle solution is replaced by a buffer solution to allow the bound vesicles to dissociate, is not always a part of the experiment, and at present, the Blivion software does not feature facilities for the analysis of this phase.

## Analysis

The purpose of the Blivion analysis protocol is to establish the two main characteristics of the binding equilibrium: the point of half saturation *Lhalf*, and the Hill coefficient *H*. *Lhalf* is related to the equilibrium dissociation constant *Kd* for the vesicle and the coated surface and is measured in response units, and the dimensionless Hill coefficient is a measure for the steepness of the binding curve at *Lhalf*. Three parameters are extracted from the response curves: the baseline response *rB* (obtained from phase 1), the loaded baseline response *rL* (obtained from phase 3), and the response *rV* generated by the vehicles bound to the coated surface at equilibrium. From these parameters the independent parameter *L*, the ligand loading, and the dependent parameter *S*, the surface coverage, are computed as follows:

The surface coverage *S* (measured in response units) and the fractional saturation *f* (dimensionless) as a function of the ligand loading *L* are empirically described by the following equations:

The three parameters, the maximum coverage level *Smax*, the half saturation value *Lhalf*, and the Hill coefficient *H* are obtained by a non-linear least-squares fit of the values for the saturation *S* as a function of the ligand loading *L*.

### Determination of the ligand loading *L* and the surface saturation *S*

In a typical experiment, the baselines obtained from phases 1 (Baseline) and 3 (Loaded) are time invariant, and *rB* and *rL* are therefore computed as the average response value over the whole or a part of the respective experimental phases. Binding of the vesicles to the coated surface occurs in phase 4 (Association), and whilst a significant approach to equilibrium should be reflected in the curves, a full equilibrium may not have been reached when this phase is terminated. The approach to equilibrium is multiphasic, but the later stages of the equilibration process are usually sufficiently accurately be described by a double exponential function:

Here, *r*(*t*) is the response as a function of time, *A1* and *A2* are the amplitudes, and *τ1* and *τ2* are the lifetimes of the two phases, whereas *A0* is the amplitude at *t* = ∞. Thus, a good fit of the later stages of the association curve to a two-exponential function predicts *r*v , the response at equilibrium. In the analysis performed by Blivion, the value of *A0*, obtained from the best double exponential fit to the later stages of the association phase is taken as the value for *rv* in the computation of the saturation value *S*.

## Using Blivion

### Input and data reduction

Pressing the Open data button on the toolbar or in the File menu opens a dialog that allows the user to select and open a data file. Blivion takes files in the comma-separated values (csv) format produced by the Octet software. The files must contain an even number of columns: two columns for each sensor, with time data in the first (odd-numbered) columns and the response data in the second (even-numbered) columns. All columns must have the same length. The odd-numbered entries in the first row must contain a unique header (name) for each pair of columns. As long as these conditions are met, Blivion will read and use the data in any even number of columns and any number of rows.

An example of the correct input format is given below, in the required csv format and, for clarity, in the tabular form in which it appears in Excel.

|  |
| --- |
| A1,,B1,,C1,,D1,,E1,,F1,,G1,,H1,  0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0  0.2,0.00302408,0.2,0.00159839,0.2,0.00482793,0.2,0.0020054,0.2,0.00079456,0.2,-0.0001861,0.2,0.00364844,0.2,-0.00212121  0.41,0.00912184,0.41,0.00090049,0.41,0.01002094,0.41,0.00778636,0.41,0.00264883,0.41,0.00516622,0.41,0.00700696,0.41,0.00322831  0.59,0.0037998,0.59,0.00090498,0.59,0.01113999,0.59,0.00485744,0.59,0.00486054,0.59,0.00474878,0.59,0.00493781,0.59,0.00096096  0.81,0.0104134,0.81,0.00231547,0.81,0.01179713,0.81,0.0105086,0.81,0.00508588,0.81,0.00608219,0.81,0.01026336,0.81,0.00167301 |

Figure 1. Cvs input data format accepted by Blivion opened in a plain text editor..

|  |
| --- |
| Figure 2. Csv input data accepted by Blivion opened in Excel. |

The first odd-numbered column (the column with the header *A1* in the above example) is taken to contain the time data; the other odd-numbered columns are assumed to contain exactly the same data as the first. Blivion moves column headers one cell to the right, gives the first column the header *time*, and discards all odd-numbered columns apart from the first. Because the number of rows generated in a typical experiment is quite large, Blivion reduces the number of rows by a given factor *m* (*m* is currently fixed at 5) by discarding all rows whose number is not a multiple of *m*.

### Determination of *rL*, *rB*, and *rv*

Figure 3 contains a screenshot of Blivion in action, immediately after opening a typical dataset. Notice that there are two toolbars, one above the main graphical area and one below it. The lower toolbar is used to manipulate and save the graphics (zoom, pan, *etc*); the upper toolbar contains the commands to set up the analysis and save the results. At the initial stage of the analysis, only the Open data and Baseline buttons in the upper toolbar are enabled. Pressing the Open data button and selecting a new dataset at any time in the analysis will discard the ongoing analysis without saving the results and start a new one with the new data.

|  |
| --- |
| Figure 3. Screenshot of Blivion immediately after opening a data file. |

Baseline. To determine the average baseline values (*rB*), press the button labelled Baseline, then place the cursor inside the panel using the mouse, press and hold the left mouse button and swipe to the left or right. While the right mouse button is pressed, a red rectangle between the top and the bottom of the main panel indicates and selects the time period swept. On releasing the right mouse button, the average over the time span swept is computed for each curve and assigned to the baseline value for that curve. Blivion will show the baseline values in the main, bottom left, panel as broken black lines, and the deviation of the individual curves from the average in the top left panel. Selection of the time period over which to compute the baseline values is up to the user, and can be anywhere in the graph, but a sensible sweep would obviously cover all or part of phase 1 and no other part of the curves, as shown in Figure 4 and Figure 5. After the first baseline sweep, the Loaded button becomes enabled, allowing the user to continue and obtain values for phase 3.

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
| Figure 4. Determining the baseline values by pressing the Baseline button and swiping the baseline area. | Figure 5. Result of baseline sweep. Average baseline values indicated by broken black lines in the bottom left graphical panel; deviations for average in the top left panel. The button labelled Loaded is now enabled. |

Loaded baseline. The average loaded baseline responses (*rL*) are obtained in the same way as the original baseline values, by pressing the Loaded button on the upper toolbar, and sweeping an area with the right mouse button pressed. Upon release of the right mouse button, the average over the swept area is computed and assigned to the average loaded baseline response. Once the values for *rB* and *rL* have been established, values for the ligand loading are computed and shown in the top right panel, and at this stage the button labelled Association has become enabled.

|  |  |
| --- | --- |
| Figure 6. Establishing the loaded baseline values by pressing the Loaded button and sweeping the loaded baseline area (phase 3). | Figure 7. Result of the loaded baseline sweep. The loaded baseline averages are indicated by broken black lines in the bottom left panel; the deviations from average in the top left panel, and the ligand loading levels in the top right panel. The button Association is now enabled. |

#### Vesicle association plateau.

During the association phase equilibrium (a plateau) is not normally reached. A reasonable estimate of the position of the plateaus can be made by performing a non-linear least-squares fit of the later stages of phase 4 with a double exponential function. Blivion automatically carries out such a fit for each curve after the Association button has been pressed and an area of the graph has been swept in the way outlined above (Figure 8). Figure 9 shows the result of the association phase sweep, with the best fit to each trace indicated as broken black lines in the bottom left panel and the deviations of the observed data from the best fit in the top left panel. Moreover, the bottom right panel now shows the saturation levels (*S*) as a function of the ligand loading (*L*) as well as the best non-linear least-squares fit of these derived data to a variant of the Hill equation (see above). The best-fit values for the *Smax*, *Lhalf* and *H* parameters are indicated inside the bottom right panel. The residuals (deviations of the observed data from the fitted curve) are plotted in the middle panel on the right.

Sweeping to determine the association level also enables the Save results button, to indicate that the results of the analysis may now be saved to a cvs file. Throughout the analysis process, the tables under the Results tab show the results obtained so far in numerical form, as shown in Figure 10.

Once they have been enabled, the Baseline, Loaded and Association toolbar buttons stay enabled throughout the analysis process, and only become disabled when a new data set is opened. If a particular button is enabled, it may be used to indicate the beginning of a new sweep. At each stage of the analysis, the results of the last sweep remain in memory. After each sweep, all applicable calculations are carried out anew.

|  |  |
| --- | --- |
| Figure 8. Determining the association plateaus for each curve by pressing the Association button and sweeping the later stages of the association phase (phase 4). | Figure 9. Result of the association phase sweep. The best least-squares fit to the individual traces is indicated with broken black lines in the bottom left panel; the deviations of the observed data from the fitted curves is in the top left panel, and a least-squares fit of the estimated saturation levels as a function of the ligand loading is shown in the bottom right panel. |

|  |
| --- |
| Figure 10. Results obtained from the analysis are available in numerical form throughout the analysis process under the Results tab. The bottom right panel contains the computed fractional saturation (y) at small x intervals (depicted as a black line in the bottom right panel under the Plots tab). |

### Critical assessment of the results

Blivion allows the user considerable freedom in choosing the most appropriate areas for the determination of the baseline values *rB* and *rL*, and the association plateau *rV*. As a consequence, users may inadvertently select areas for the determination of these values that are less suitable. The algorithms used for the non-linear least-squares fitting of the association traces and the surface saturation data are quite robust, and will under most conditions produce a best fit of the data with the designated function.

It is the responsibility of the user to make sure that the baseline and loaded baseline averages and the estimated association plateaus pertain to the correct phase of the trace; Blivion does not check or provide warnings. Furthermore, users are advised to check that the estimated plateaus for the association phase are realistic and representative for the available data. Figure 11 shows the results of a typical flawed analysis, in which the residuals for the best fit of three of the traces with a double exponential function are non-randomly distributed along the time axis. This indicates that the observed data over this time range are not sufficiently accurately described by a double exponential function, and probably reflect the remnants of a faster process. As a result, the saturation level at equilibrium is underestimated, which, in turn, is reflected in an underestimate of *Smax*, the maximum saturation level under the given experimental conditions. Although the characteristics of such a faster phase could, in principle, be revealed by a least-squares fit to a triple or higher exponential function, in practice the fitted parameters are too highly correlated for such a fit to yield a meaningful and well-defined estimate of the plateau values. The solution is, therefore, to avoid including faster phases by limiting the range to the fit to the later stages of the association process.

|  |  |
| --- | --- |
| Figure 11. Example of a flawed analysis: the association plateaus for the red, green and blue traces are underestimated. | Close-up of the orange, red, green, and blue traces and their best fit (broken black lines). |

### Output to csv

Once the analysis has yielded a satisfactory result, the observed or estimated values for *rB*, *rL*, *rV*, the values for *L* and *S* computed from the observations, the best-fit parameter values for the fit of *S* to the variant of the Hill equation above, and the numerical values representing the solid black line in the bottom right panel of the user interface of Blivion may be saved as a cvs file that can be opened in Excel, by pressing the Save results button in the top toolbar. An image of the graphics under the Plots tab may be saved in various graphical formats (png, jpeg, pdf, and several more) by pressing the Save the figure button in the bottom toolbar.

|  |
| --- |
|  |
| Figure 12. Blivion output: analysis results saved in cvs format as read by Excel (top panel), and as an image (various graphical formats available; bottom panel). |

## Technical details

Blivion was conceived by Stephen Martin and implemented in python 3.5, using python’s pandas, matplotlib and scipy libraries for the data manipulation, visualisation and analysis. The graphical user interface was implemented in Qt5 using the pyqt5 bindings, and the executable version was created using PyInstaller.