Introduction

The goal of the BioSAXS experiment with number MX-2102 was to compare the general shape and flexibility of several *de novo* designed polypeptides. The project involves a novel strategy to push forward the field of protein design, where a computational approach is used to generate large amounts of protein scaffolds to sample the right protein conformation for supporting a functional motif. Before advancing into methods for high-resolution structure determination, quick, low-resolution structural tools are in need to high-throughput screen the competent constructs. Combined with other biophysical characterization data, for example CD (circular dichroism) and MALLS (multi-angle laser light scattering), we propose SAXS to support the practicability of the computationally *de novo* designed proteins. We therefore aimed to collect data on several of these designed proteins belonging to the same topological category.

Data acquisition

The data were recorded at the ESRF BM29 over a q range of 0.25-0.5 $Å^{-1}$ and beam wavelength of 0.992 Å. Ten designed proteins were analysed in a buffer which was composed of PBS and 10 % glycerol. Nine of the proteins were well-behaved and the higher-throughput batch mode method was used. Protein samples were exposed under flow passing through a capillary at a flow rate of 10 or 8 $\mu L \cdot s^{-1}$ with 10 individual frames collected and averaged to give the final scattering profile. The dilution series was done diluting the protein in the buffer mentioned above. The blank for each measurement was the same buffer. Scattering from this black alone was collected before and after each sample to correct for fluctuations in beam intensity and to clean the capillary between samples. The temperature was maintained always at 20° C and the viscosity was set to low. We used 100% of the beam transmission. The concentration series ranged from 10.5 to 2.5 mg/ml. Data of the one protein construct showing signs of aggregation were measured using the HPLC mode (Superdex 75 column (24 ml, GE Healthcare)) at a flow rate of 0.5 ml/min in the same buffer. 100 ul of the sample were injected at a concentration of 15 mg/ml. During the chromatography run, data were collected for 50 min, with one frame per second, originating 3000 frames per run.

Data analysis

Most of the proteins behaved well and gave good scattering elution profiles that were analysed using SCÅTTER [1] and the ATSAS package [2].

Results and Conclusions

During this shift, we have obtained SAXS data on ten different constructs. In **Figures 1** and **2** and **Table 1** we show some examples of the data sets collected.

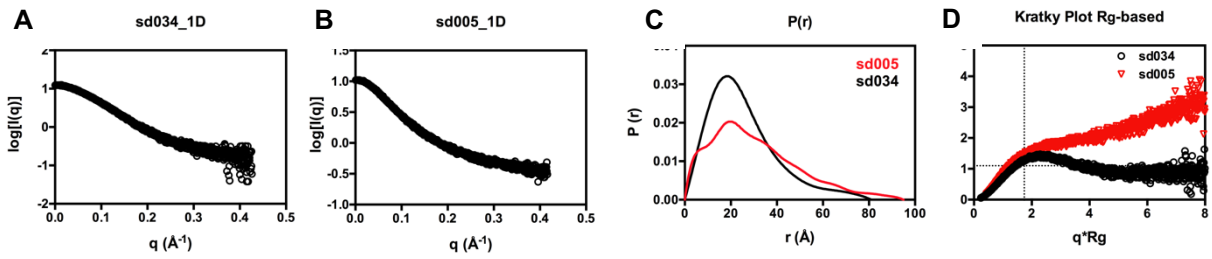

Figure 1: SAXS analysis of the sd034 and sd005 constructs. **A** & **B**) Scattering intensity plots. **C**) P(r), pair-distance distribution, function. **C**: Kratky plot showing the difference in flexibility between

constructs. Both constructs show similar CD spectra but adapt different conformational space as shown by the P(r) and the Kratky plots, in accordance with the results from NMR spectroscopy.

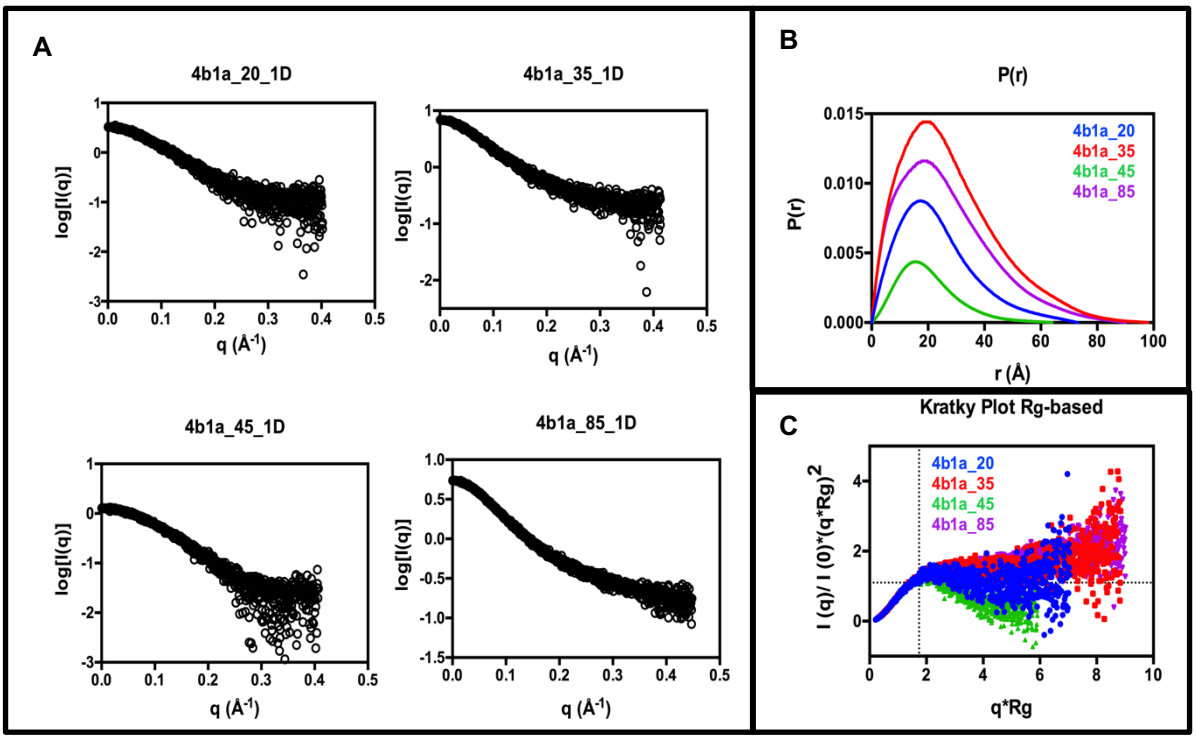

Figure 2: SAXS analysis of four 4b1a constructs belonging to the same topological category. **A**: SAXS scattering curves. **B**: P(r) function. **C**: Kratky plot.

Construct	Data range (Å^{-1})	R_g Guinier (Å)	D_{max} (Å)	Theoretical MW (kDa)
4b1a_20	0.45-40.2	17.71 ± 0.22	73	7.8
4b1a_35	0.66-41.3	21.42 ± 0.23	98	7.2
4b1a_45	0.47-40.1	15.73 ± 0.08	64	6.9
4b1a_85	0.51-44.8	20.41 ± 0.11	90	7.1

Table 1: SAXS data analysis parameters. 4b1a constructs were *de novo* designed as functional proteins to present immunological epitopes. Design 4b1a_45 showed the most compact propensity of all variants, and it will be the candidate for more advanced structural characterization.

With these data in hand, we expect to compare the radius of gyration (R_g), maximum dimension (D_{max}) and Kratky plots of the different designed proteins within the same topological space in order to distinguish the rigid constructs from the partially folded ones. This will serve as a strong experimental evidence to feedback the computational algorithm toward designing folded protein *in silico*. Using SAXS together with other complementary techniques such as CD, DLS and SEC-MALLS, we will achieve an integrated computational and experimental method that enables a high-throughput pipeline for validating and testing the large number of protein designs.

References

1. ScÅtter - bioisis.net - Rambo RP and Tainer JA. Biopolymers (2011):p. 559-571.
2. Franke, D *et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions.* J. Appl. Cryst. (2017) **50**, 1212-1225.