The optimization of in vitro high-throughput chemical lysis of *Escherichia coli*. Application to ACP domain of the polyketide synthase ppsC from *Mycobacterium tuberculosis*

Pawel Listwan · Jean-Denis Pédelacq · Meghan Lockard · Carolyn Bell · Thomas C. Terwilliger · Geoffrey S. Waldo

Abstract Protein production in *Escherichia coli* involves high-level expression in a culture, followed by harvesting of the cells and finally their disruption, or lysis, to release the expressed proteins. We compare three high-throughput chemical lysis methods to sonication, using a robotic platform and methodologies developed in our laboratory [1]. Under the same expression conditions, all lysis methods varied in the degree of released soluble proteins. With a set of 96 test proteins, we used our split GFP to quantify the soluble and insoluble protein fractions after lysis. Both the amount of soluble protein and the percentage recovered in the soluble fraction using SoluLyse® were well correlated with sonication. Two other methods, Bugbuster® and lysozyme, did not correlate well with sonication. Considering the effects of lysis methods on protein solubility is especially important when accurate protein solubility measurements are needed, for example, when testing adjuvants, growth media, temperature, or when establishing the effects of truncation or sequence variation on protein stability.

P. Listwan (✉) · J.-D. Pédelacq · M. Lockard · C. Bell · T. C. Terwilliger · G. S. Waldo
Bioscience Division, MS-M888, Los Alamos National Laboratory, Bikini Atoll Rd, SM30, Los Alamos, NM 87545, USA
e-mail: plistwan@lanl.gov

J.-D. Pédelacq
Institut de Pharmacologie et de Biologie Structurale (IPBS), CNRS, Toulouse, France

J.-D. Pédelacq
UPS, IPBS, Université de Toulouse, Toulouse, France

Abbreviations

GFP	Green fluorescent protein
SAMI	SAGIAN™ automated method interface
HTP	High-throughput

Introduction

One of the benefits afforded by the International Structural Genomics Organization initiative (ISGO) is the development of new automated, high-throughput (HTP) protein production technologies. Soluble protein production for structural and functional determination is essential, yet is also one of the most difficult parts of the sequence-to-structure pipeline. The Protein Structure Initiative (PSI) in United States, the Structural Proteomics in Europe (SPINE), and other partner programs around the world have established a variety of automated methodologies for processing a large number of protein constructs [2–8]. Despite years of research and development in protein production, automation, and HTP technology, no single cell disruption methodology exists that satisfies the needs of all structural genomics laboratories.

Automated, HTP protein production as the instrument for structure determination is a complex, multistep process that requires the optimization of each individual task. *Escherichia coli* continues to be a popular host for protein expression despite the large proportion of the recombinant proteins that often accumulate as aggregates or inclusion bodies [9]. Changing the expression host to insect cells, baculovirus, cell-free expression or mammalian expression often presents additional problems. For this reason many laboratories and commercial institutions direct great efforts

to improve bacterial expression strains, vector systems, and other factors that improve recombinant protein expression and solubility.

One of the most crucial steps to be optimized in the protein production process is bacterial cell lysis. Although bacterial cell lysis does not influence protein expression, it can have an effect on protein solubility by affecting the physicochemical properties of the protein. Conventional biochemistry laboratories working with a few protein targets can test and optimize many lysis methods. These include techniques such as mechanical cell disruption—e.g. sonication, french press, and freeze-thaw, and chemical lysis using different buffer composition, lysozyme, or commercially available detergent reagents. Cell lysis can also include a combination of the mechanical and chemical lysis, e.g. lysozyme with freeze-thaw cycles. The preferred method, or "gold standard", for bacterial lysis on the small or standard laboratory scale production is sonication. It relies on the mechanical disruption of the bacterial cell wall. The expressed protein is not affected by any solubilizing lysis agents, like detergents, that can affect solubility or stability [10, 11]. On the other hand, when hundreds or thousands of different proteins, truncation, or sequence variants are screened, only a few lysis methods can be reasonably employed. Sonication becomes more problematic when hundreds of proteins need to be released from the bacteria using automated, HTP liquid handling platforms. Although there are HTP sonicators available on the market, e.g. SonicMan (MatriCal, Spokane, WA), most structural genomics liquid handling platforms were established before the availability of the HTP sonicators. Additionally, the selection of high-throughput sonicators is still very limited and costly [consequently are often difficult to integrate with current laboratory setups]. For this reason many HTP laboratories choose to optimize lysis conditions by chemical means.

As a member of the Integrated Center for Structure and Function Innovation (ISFI), part of the PSI Specialized Center Program, we are focused on developing methods that overcome bottlenecks in soluble protein production and protein crystallization. The split-GFP technology developed in this laboratory [12–16] has recently been used to develop an automated, HTP solubility screening assay, allowing us to process and screen thousands of protein constructs for solubility in a few days [1]. Briefly, split GFP technology uses highly engineered, self-complementing GFP fragments originally derived from "superfolder" GFP: a 15 amino acid GFP "tagging" fragment—strand 11 (S11 or GFP 11) and a GFP 1-10 "detector" fragment. The GFP S11 fragment is fused to the C-terminus of the protein of interest in a pTET plasmid. GFP 1-10 is separately expressed in a pET plasmid. The S11 fragment is available for complementation by the GFP 1-10 fragment only if the protein of interest is stable and soluble. This spontaneous complementation leads to formation of the fluorescent GFP beta-barrel.

Screening terminal deletion libraries with the split GFP in order to identify compact, soluble domains can facilitate structural study of large, multidomain proteins. The measured solubility and sequenced ends of each fragment from the library are mapped onto the protein's sequence, providing a comprehensive roadmap of soluble expression as a function of $5'$ and $3'$ construct ends.

The objective of library screening is to evaluate the intrinsic solubility and stability of each member. Even single amino acid extensions or deletions at either end of the protein can profoundly affect expression. It is important to control the effects of chemical lysis on protein stability in order to reliably and accurately measure the effects of amino acid mutations or terminal deletions.

To help accomplish this goal, we have tested several lysis reagents with a library of protein constructs and have compared the solution chemical lysis methods to sonication. Here, we compare the solubility data obtained from a library of different size constructs of the ppsC's gene, originating from ACP domain and spanning up to the two adjacent domains; KR and ER. The lysis methods used include lysozyme, freeze-thaw cycles, Bugbuster, SoluLyse, and sonication. The goal of this experiment was to identify a chemical lysis method for our automated, HTP solubility assays that gives results that best match manual low throughput sonication.

Materials and methods

Robotics integration

Our integrated, high-throughput robotic system has previously been described [1]. Briefly, it includes a Biomek FX liquid handling robot, an ORCA® arm, a DTX plate reader equipped with filters allowing measurement of both absorbance and fluorescence (Beckman-Coulter, Fullerton, CA), a Cytomat 24 Hotel, Cytomat 2C incubators (ThermoFisher Scientific, Waltham, MA) and a Rotanta 46 ESC centrifuge (Hettich AG, Tuttlingen, Germany). For fluorescence imaging we used either an Illumatool lighting system LT-9500 (Lightools Research, Encinitas, CA) or Run Time Data Viewer, simulation software, version 3.0.0.9 (Beckman-Coulter, Fullerton, CA). Cultures were grown overnight in Innova 4230 refrigerated incubator shakers (New Brunswick Scientific, Edison, NJ). Sonication of the 96-well plates was performed manually using a Sonicator—ultrasonic processor XL20-

20 (Misonix Inc., Farmingdale, NY). Biomek FX methods are designed using Biomek Software, version 3.2. Integration of all robotics components was controlled by SAMI Method Editor, version 3.5 (Beckman Coulter, Fullerton, CA).

Expression library generation

For the test of the different lysis methods using a liquid handling platform we used a library of 96 protein constructs representing the Acyl carrier protein (ACP) domain of the *Mycobacterium tuberculosis* Polyketide Synthase (ppsC) (Genbank accession number: CAB06099.1). The samples were prepared as previously described [12]. Following the library preparation, all plasmids were expressed in *Escherichia coli* BL21 (DE3) strain (Stratagene, La Jolla, CA). Overnight culture growth from the library's glycerol stock was performed in 175 µl Luria-Bertani (LB) media supplemented with 7.5% glycerol and selective antibiotics spectinomycin (75 µg/ml) and kanamycin (35 µg/ml), standing at 32°C for 16 h. Ten microlitres was used to inoculate 1 ml of LB media supplemented with antibiotics in a 96 deep-well plate. Following 2 h outgrowth at 32°C, 350 rpm in Innova 4230 refrigerated incubator shaker, protein expression was induced by anhydrotetracycline. Cultures were grown for additional 2 h at 32°C, and then quenched using chloramphenicol (Sigma–Aldrich, St. Louis, MO). Equal volumes of expressed bacterial cultures (175 µl) were transferred to 4 × 96-well microtitre plates and centrifuged at 4,000 rpm for 15 min. The supernatant was removed from all plates. Bacterial pellets were dried and stored at −80°C before lysis. Buffers used throughout the experimental procedure included either TNG buffer (100 mM Tris–HCl (pH 7.4), 150 mM NaCl, 10% glycerol) or TN buffer (100 mM Tris–HCl (pH 7.4), 150 mM NaCl). GFP 1-10 detector fragment reagent was prepared as described previously [12].

Lysis methods

Chemical cell lysis was performed using 120 µl lysis solution containing either lysozyme (Sigma–Aldrich, St. Louis, MO) and 2 freeze-thaw cycles at −80°C, SoluLyse® in Tris buffer (Genlantis, San Diego, CA) or Bugbuster® protein extraction reagent (Novagen, EMD Chemicals Inc., San Diego, CA). For all commercial lysis reagents, and the lysozyme, manufacturer's protocols were followed with the addition of the Benzonase® (Sigma–Aldrich, St. Louis, MO). For manual sonication, 120 µl of TN buffer was added to of 96-well plates containing centrifuged cells that were then sonicated using ultrasonic processor XL20-20, 3 × 90 s, 50% cycle.

GFP 1-10 complementation and fluorescence data measurement

Following lysis, soluble and insoluble fractions were separated by centrifugation in Rotanta 46 RSC integrated centrifuge at 4,000 rpm for 20 min, 4°C. GFP 1-10 complementation was achieved by the addition of 190 µl of GFP 1-10 reagent into 40 µl of the soluble fraction and 190 µl of GFP 1-10 reagent into 10 µl of the solubilized pellet fraction (previously denatured using 60 µl 9 M urea in TNG). The split GFP can be used to measure as little as 0.2 pmol of protein in as little as 30 min using kinetics [14]. However, we chose to take advantage of the stability of the reconstituted GFP [12]. The final fluorescence value was measured after 24 h. This eliminated possible time dependence of the readings and simplified calibration and measurement of many samples. GFP 1-10 complemented plates were incubated overnight at 4°C and the final fluorescence was measured using DTX reader. For fluorescence images, plates were illuminated using an Illumatool lighting system LT-9500. For the quantification of the soluble protein fraction, a set of eight different concentrations of the soluble GFP S11 tagged control protein, sulfite reductase, was obtained by serial dilution and used to generate a calibration curve as previously described [1]. The estimated expression yield (estimated mg/l) from the test protein cultures were each calculated using a calibration curve and the final fluorescence and calculated molecular weights of the protein fragments (all fragments were sequenced) as previously described [14].

Finally, the left and right ends of each construct and the measured soluble protein, insoluble protein, and fraction soluble were each mapped onto the genomic DNA sequence of the complete ppsC gene to visualize compact soluble domain boundaries.

Results

In this comparative study, we tested three different chemical lysis methods that can easily be automated and integrated into any HTP liquid handling robotic platform. Sonication was used as the standard of the comparison. A single 96 well plate containing picks from the ACP library was grown and used to inoculate four replicates for induction. All four methods were tested on expressed bacterial cell pellets originating from these replicates, under identical growth and expression conditions. For each plate, we used our split GFP to assay the soluble and insoluble protein fractions after disruption and centrifugation. Complementation with exogenous GFP 1-10 fragment resulted in a range of fluorescence depending on the solubility level of the tested protein constructs. Figure 1
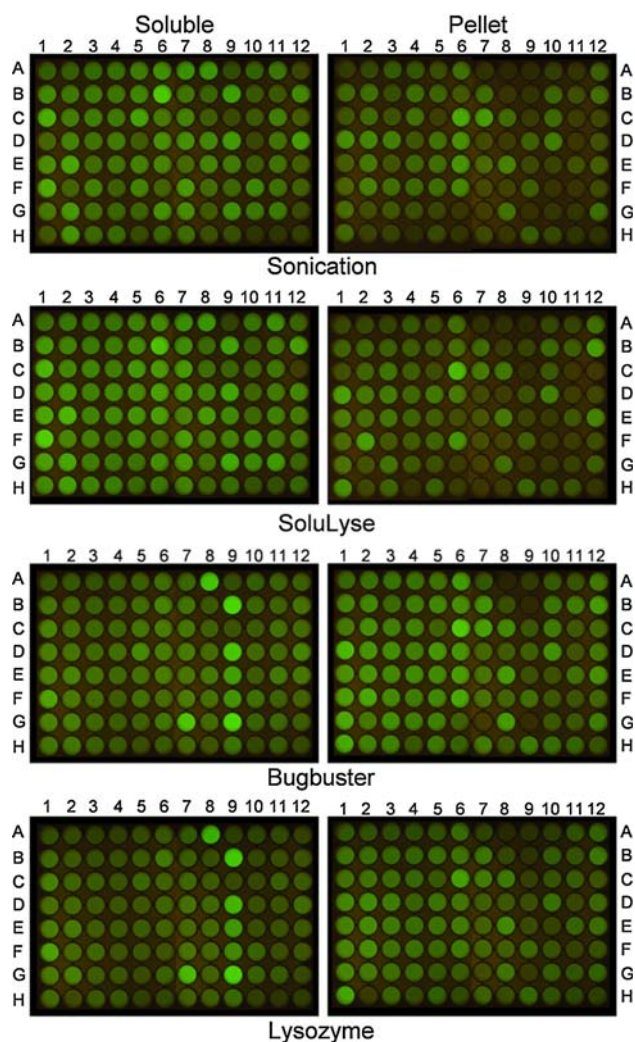
Fig. 1 Fluorescence data for 96 ACP domain ppsC library constructs obtained from four different lysis methods. Both soluble and pellet fractions are shown. Each well on four plates represents a single protein construct that was expressed using same conditions but lysed with one of the four lysis methods

represents the fluorescence images of both soluble and pellet fractions for all four lysis methods used in this study after their complementation with GFP 1-10. Each well on a given plate represents a single, unique ACP domain construct. The fluorescence signal for the same well position on different plates is indicative of the variation between the different lysis methods for the same construct. The sum of the fluorescence value for each soluble fraction and its corresponding pellet fraction represents the constructs total fluorescence. Because the absorbance value for each well is the same between four plates (data not shown), the difference between fluorescence values, and therefore relative solubilities, is reflective of the lysis method. Figure 2 shows the correlation between different chemical lysis methods and with

sonication. The fluorescence data obtained for one lysis method was plotted against the data for the second lysis method. Figure 2a shows graphs and the correlation between sonication and different chemical lysis methods. The highest correlation value is for the SoluLyse reagent (correlation coefficient of 0.74). Both lysozyme and Bugbuster methods poorly correlate with sonication (correlation coefficients of 0.33 and 0.30, respectively). Figure 2b shows correlation between three chemical lysis methods. Lysozyme and Bugbuster methods have the highest correlation with each other (correlation coefficient of 0.97). Table 1 combines the total expression and the percentage of the released soluble protein using four lysis methods. The position on the plate, length of the construct, molecular weight, total expression (mg/l), and the percentage of the soluble fraction is presented. Figure 3 maps the ACP domain constructs onto the ppsC sequence used in this study. The six colors represent fragments with increasing solubility percentage (0–17% red, 17–34% orange, 34–50% yellow, 50–67% green, 67–84% dark blue and 84–100% light blue (expressed as the percentage of the total protein expression). The pattern of solubility in panels A and B of Fig. 3 are quite similar, as are those in panels C and D. This is consistent with the results shown in Fig. 2, where the sonication and SoluLyse® methods are correlated. Neither the BugBuster® nor lysozyme methods correlate with sonication, suggesting a protein dependent bias of solubility or lysis efficiency relative to sonication. The strong correlation of Bug-Buster® with lysis by freeze thaw suggests a similar mode of bias. Table 1 summarizes the calculations of the total expression yield and the success of the lysis method as a function of percentage of the soluble fraction released from the bacterial cells. Despite the fact that all plates contained the same amount of protein expressed before the lysis, there is a variation in total expression yield between the methods as detected by the complementation with GFP 1-10. These results suggest that chemical reagents not only rupture the bacterial cells but also have an effect on the proteins' physicochemical properties by affecting the exposure of the GFP S11 tag on the construct to complementation by GFP 1-10 and thus final fluorescence. The results clearly indicate that none of the chemical methods tested are identical to sonication. SoluLyse® reagent was found to be the most similar to sonication. Lysozyme and Bugbuster® lysis showed poor correlation with sonication under the conditions tested. However, these methods produce results similar to each other, especially for smaller protein constructs. The noticeable outliers are found in Fig. 1, positions A8, B9, D9, G7 and G9. The same constructs can be seen in Fig. 3, as the only fragments 50 percent or more soluble (dark blue and light blue colors). These are the

Table 1 Comparison of expression yield and lysis method effectiveness using four lysis methods

R:C	aa	MW	Sonication		Solulyse		Bugbuster		Lysozyme	
			Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)
A1	436	536,81	34	51	50	52	60	27	53	29
B1	442	544,40	61	47	68	56	75	28	61	29
C1	411	506,04	55	70	59	76	59	37	53	39
D1	471	583,40	67	20	101	31	119	18	64	24
E1	420	516,74	36	57	63	50	64	30	52	31
F1	412	507,21	47	63	70	69	81	37	62	43
G1	439	540,45	42	38	61	51	89	22	57	28
H1	459	568,04	42	25	89	32	113	18	94	15
A2	442	544,40	49	43	65	56	86	24	56	30
B2	438	539,13	69	37	69	46	92	23	60	24
C2	425	523,83	55	42	58	49	83	24	55	27
D2	439	540,45	74	32	77	47	89	24	71	21
E2	423	520,43	71	50	76	62	82	28	69	26
G2	412	507,21	45	68	56	71	61	36	56	36
H2	403	496,44	47	71	58	74	76	31	17	44
A3	435	535,62	73	40	66	54	95	22	56	27
B3	435	535,62	58	54	73	45	71	30	62	24
C3	442	544,40	53	53	86	40	84	25	60	26
D3	425	523,83	77	27	82	36	90	21	66	20
E3	426	524,72	41	45	73	40	86	25	59	23
F3	434	534,73	64	40	53	51	84	23	53	29
G3	426	524,72	46	43	74	38	93	20	52	25
H3	433	533,08	43	52	81	47	94	22	54	27
A4	436	536,81	68	46	73	46	88	24	56	27
B4	439	540,45	75	37	68	47	86	24	63	24
D4	443	545,86	53	50	73	48	81	28	53	27
E4	435	535,62	61	39	71	50	95	22	60	24
G4	439	540,45	50	44	65	48	92	21	62	24
H4	412	507,21	38	70	54	75	61	33	46	38
A5	419	515,69	54	56	58	55	70	27	50	31
B5	434	534,73	59	49	59	52	80	24	63	24
C5	411	506,04	47	79	49	68	46	43	44	40
D5	436	536,93	49	24	74	40	103	29	60	22
E5	442	544,40	63	41	63	56	84	24	64	24
F5	435	535,62	67	42	69	51	87	24	54	25
A6	385	473,35	62	47	63	43	87	21	54	23
B6	412	507,21	69	75	78	64	71	31	59	34
C6	459	568,04	91	20	128	24	137	14	92	15
D6	439	540,45	74	33	67	47	74	26	53	26
E6	439	540,45	77	28	47	71	83	24	59	24
F6	567	701,78	81	20	104	24	133	16	81	18
G6	411	506,04	42	78	44	76	53	43	47	36
H6	412	507,21	35	59	48	65	61	29	42	32
A7	411	506,04	52	77	54	77	53	42	42	44
B7	439	540,45	55	36	67	46	88	24	59	25
C7	425	523,83	81	26	69	44	86	21	54	25
D7	404	497,45	32	68	36	72	41	42	33	43

Table 1 continued

R:C	aa	MW	Sonication		Solulyse		Bugbuster		Lysozyme	
			Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)	Total (mg/l)	Sol (%)
E7	365	451,10	29	32	27	58	38	33	27	37
F7	395	487,33	29	74	33	72	40	38	32	39
G7	160	198,36	11	78	13	75	21	89	18	76
A8	152	189,75	20	87	25	79	38	94	33	87
B8	401	493,61	24	64	32	58	35	38	29	37
D8	404	497,45	31	74	33	72	38	42	33	40
E8	436	536,81	64	34	71	46	87	23	63	21
F8	358	443,45	19	60	26	66	35	38	26	36
G8	486	600,14	49	24	57	35	97	19	50	25
H8	298	367,89	14	52	19	57	27	32	20	41
A9	288	358,26	19	60	18	57	29	45	23	42
B9	152	189,75	22	82	28	73	42	95	38	88
C9	454	559,97	24	43	26	74	43	32	28	38
D9	152	189,75	25	57	25	69	39	70	28	72
G9	152	189,75	15	84	23	84	37	94	31	87
A10	443	545,86	46	51	65	52	68	35	48	38
B10	465	575,63	58	27	67	29	82	27	56	25
C10	488	602,36	60	25	52	40	78	21	46	27
D10	395	486,44	56	14	70	24	85	17	55	20
E10	368	454,77	25	47	26	63	33	39	25	38
F10	411	506,04	40	76	48	78	52	46	43	39
G10	412	507,21	40	70	51	75	54	38	52	33
H10	465	575,63	40	21	59	35	86	19	50	22
A11	411	506,04	36	69	58	68	46	41	49	31
C11	406	499,75	30	69	26	73	38	45	43	34
D11	112	141,03	7	70	8	69	9	45	8	40
E11	406	499,71	36	48	45	53	58	38	44	25
F11	395	487,33	26	71	33	68	36	40	29	39
G11	410	504,73	37	71	62	74	51	42	48	35
H11	487	601,31	29	33	60	32	84	28	49	21
A12	486	600,14	39	23	60	30	65	25	54	21
B12	412	507,21	47	44	86	36	71	22	50	20
C12	452	557,19	19	50	9	42	37	35	31	36
D12	410	504,73	38	68	28	70	50	38	44	34
F12	452	557,19	27	46	36	58	42	32	35	34
G12	487	601,03	53	19	49	38	75	22	53	22

Expression yield is represented as mg/ml of total protein while lysis method effectiveness is expressed as the percentage of released soluble protein (% sol). The position on the plate (R:C), the number of amino acids (aa), and the construct molecular weight (MW) is presented for all constructs

smallest fragments covering primarily the ACP domain (fragments size range 152–160 amino acids). It is likely that these smaller proteins are relatively insensitive to the choice of lysis method. Interestingly, except for these small fragments, all of the remaining protein constructs show poor solubility (<50%) relative to sonication or SoluLyse®.

Discussion and conclusions

Structural genomics laboratories around the world are striving to develop robust expression screening methods on a small scale using liquid handling platforms that would identify highly soluble proteins amenable for scale-up production and structural analysis. In some cases, the
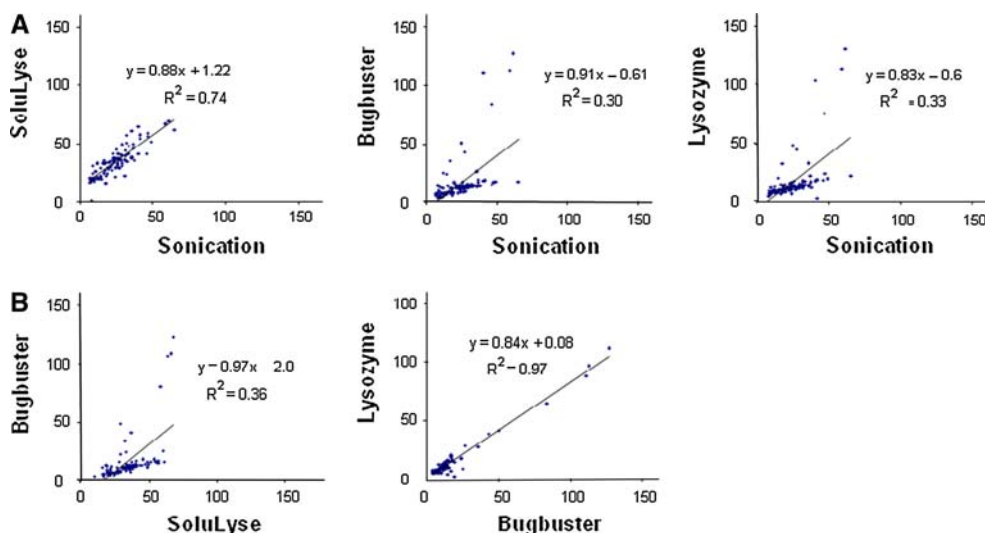
Fig. 2 a Correlation analysis between three different chemical lysis methods and sonication using the fluorescence values (all *F*-values shown are E + 06). Lysozyme, SoluLyse and Bugbuster reagents were used to chemically disrupt bacterial cells for the release of the expressed proteins. Similarly, the same constructs were lysed mechanically by sonication ("gold" standard). **b** Additionally, correlation between three chemical methods is also shown. From the plotted data SoluLyse reagent shows to be most similar to sonication. Correlation value (*R* value) for lysozyme and the Bugbuster shows that these two methods are also most similar

identification of the well behaved or totally insoluble targets is straightforward. However, most proteins expressed in *Escherichia coli* show varying degrees of solubility. Protein solubility can be affected by numerous factors, including vector design, solubility helpers, expression partners that can rescue many otherwise insoluble proteins, etc. An often neglected factor that can have a detrimental effect on the amount of released soluble protein from the bacterial cells is the lysis method. Several sources, including a comparative study of the SPINE consortium laboratories [2] and PSI protein production centers, vary widely in the use of lysis methods. No Center has systematically tested the chosen lysis method against other established approaches.

One of our laboratory's goals is to establish an automated, HTP solubility screening method using our Split-GFP technology. The method is routinely applied to full length proteins in solubility screening or to domain trapping of multi-domain proteins. In the later case, the correct solubility information is crucial in the prediction of the optimal domain boundaries, where a difference in a single amino acid can profoundly affect solubility, expression, and stability.

The purpose of this study is to benchmark alternative lysis method against sonication, in order to use the optimal method in our automated, HTP protein solubility assay, which would not require the implementation of the 96-well high throughput sonicator. We used a library of constructs covering an ACP domain region of the ppsC multidomain protein. The overall low solubility of proteins in Bugbuster® and lysozyme causes most of the larger proteins to

measure as insoluble, compressing most of the data points into a narrow region (Fig. 2). Consequently, most of the fitting power in the correlation plot between the Bugbuster® and lysozyme is largely based on a small subset of the proteins being effectively released (the smaller fragments). In contrast, most of the proteins are successfully released by sonication, without bias to size of the protein. It is also important to note that the solubility of the fragments spanning the three ppsC domains is also related to the known boundaries of the domains. Figure 3 shows that most of the fragments reach their highest solubility values when approaching the boundary of the linker region between the ER and KR domains. Larger fragments that originate in the ER domain are predominantly <50% soluble. The solubility of these fragments can be affected by the larger size of the fragment itself and/or the ER domain's incomplete or disrupted folding affecting the folding of the rest of the protein construct. Because SoluLyse® is well correlated with sonication indicates that this chemical method is acceptable for this work (Fig. 2). This can also be visualized in Fig. 3, where the most soluble Solulyse and sonication fragments indicate the boundary between the ER and KR domains.

In conclusion, the purpose of this study is to evaluate several chemical lysis methods for the release of the soluble proteins from their bacterial expression host. When bacterial cells preparations are lysed on a small scale, majority of laboratories use the sonication as the most efficient method for complete cell rupture. Moving from laboratory bench to the HTP robotics platform often creates several bottlenecks, one of them being the
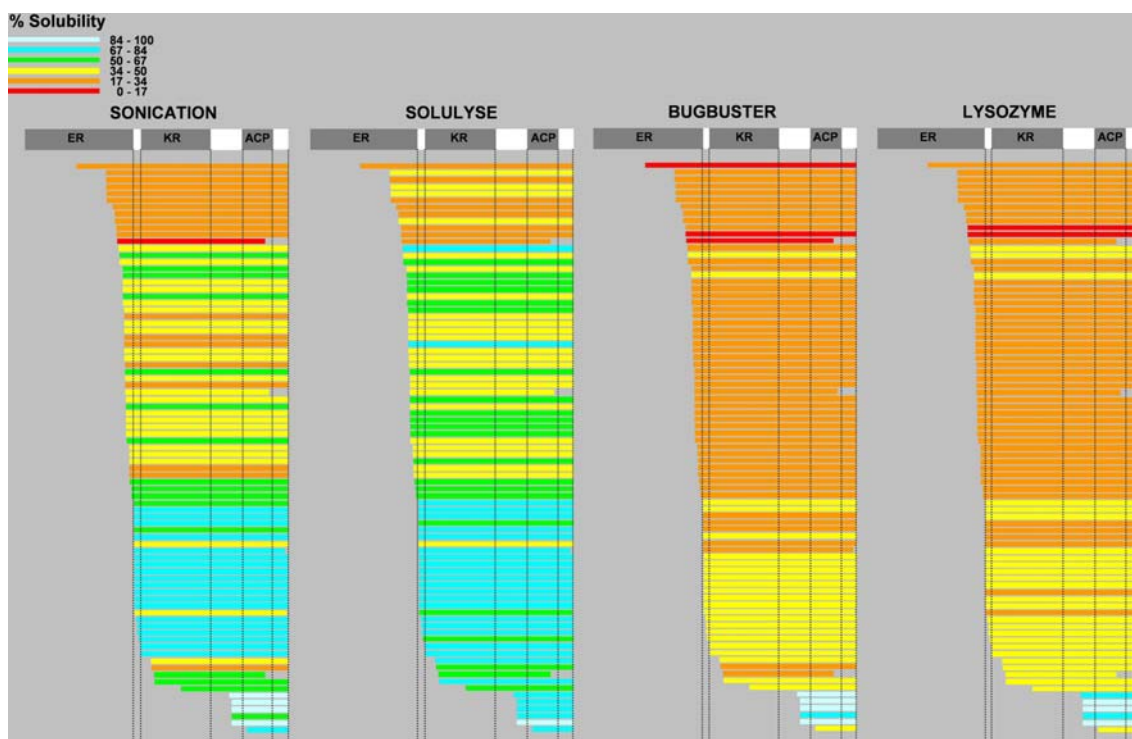
Fig. 3 Mapping of the solubility data obtained using four different lysis methods. Each construct is sequenced, assigned solubility level and mapped onto the region of the ppsC gene containing ACP domain. Colors of the *bars* represent percentage solubility as calculated by comparison of soluble and pellet fractions. Colors *red*, *orange* and *yellow* represent constructs with <50% calculated soluble fraction. Colors *green*, *dark* and *light blue* represent constructs with above 50% solubility. *Vertical dotted lines* represent the known boundaries of the ACP, KR and ER domains. To simplify the view of the map, only three—C-terminal domains of the ppsC gene are shown

availability of the high number of samples processing sonicators and/or their integration with the existing robotics platforms. For this reason, most HTP laboratories develop an in-house chemical lysis method, use commercially available chemical reagents with or without any modification, or include a manual sonication step, where the plates are moved away from the robotics platform and processed using stand alone sonicators. In our comparison study, SoluLyse® has shown to be the method with the highest correlation to sonication. Interestingly, widely used lysozyme and Bugbuster® failed in most our cases to completely release the soluble protein. In applications such as protein domain trapping, the precise solubility information is critical in predicting the boundaries of the domains. It is therefore of outmost importance that the lysis method used in such applications be non-perturbing. When sonication is inconvenient, or difficult to apply to many samples, we find SoluLyse® to be an acceptable alternative for the ACP fragment proteins as well as many other constructs we have examined to date.

References

1. Listwan P, Terwilliger TC, Waldo GS (2009) Automated, high-throughput platform for protein solubility screening using a split-GFP system. J Struct Funct Genomics 10:47–55
2. Berrow NS, Büssow K, Coutard B, Diprose J, Ekberg M, Folkers GE, Levy N, Lieu V, Owens RJ, Peleg Y, Pinaglia C, Quevillon-Cheruel S, Salim L, Scheich C, Vincentell R, Busso D (2006) Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study. Acta Crystallogr D Biol Crystallogr 62:1218–1226
3. Kim Y, Dementieva I, Zhou M, Wu R, Lezondra L, Quartey P, Joachimiak G, Korolev O, Li H, Joachimiak A (2004) Automation of protein purification for structural genomics. J Struct Funct Genomics 5:111–118
4. Mayer MR, Dailey TA, Baucom CM, Supernak JL, Grady MC, Hawk HE, Dailey HA (2004) Expression of human proteins at the southeast collaboratory for structural genomics. J Struct Funct Genomics 5:159–165
5. Vincentelli R, Canaan S, Offant J, Cambillau C, Bignon C (2005) Automated expression and solubility screening of His-tagged proteins in 96-well format. Anal Biochem 346:77–84

6. Nguyen H, Martinez B, Oganesyan N, Kim R (2004) An automated small-scale protein expression and purification screening provides beneficial information for protein production. J Struct Funct Genomics 5:23–27

7. Finley JB, Qiu SH, Luan CH, Luo M (2003) Structural genomics for caenorhabditis elegans: high throughput protein expression analysis. Protein Expr Purif 34:49–55

8. Tegel H, Steen J, Konrad A, Nikdin H, Pettersson K, Stenvall M, Tourle S, Wrethagen U, Xu LL, Yderland L, Uhlén M, Hober S, Ottosson J (2009) High-throughput protein production—lessons from scaling up from 10 to 288 recombinant proteins per week. Biotechnol J 4:51–57

9. Mitraki A, King J (1989) Protein folding intermediates and inclusion body formation. Nat Biotechnol 7:690–697

10. Feliu JH, Cubarsi R, Villaverde A (1998) Optimized release of recombinant proteins by ultracentrifugation of *E. coli* cells. Biotech Bioeng 58(5):536–540

11. Ho CW, Chew TK, Ling TC, Kamaruddin S, Tan WS, Tey BT (2006) Efficient mechanical cell disruption of *Escherichia coli* by an ultrasonicator and recovery of intracellular hepatitis B core antigen. Process Biochem 41:1829–1834

12. Cabantous S, Waldo GS (2006) In vivo and in vitro protein solubility assays using split GFP. Nat Methods 3:845–854

13. Cabantous S, Terwilliger TC, Waldo GS (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat Biotechnol 23:102–107

14. Cabantous S, Pédelacq JD, Mark B, Naranjo C, Terwilliger TC, Waldo GC (2005) Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from *Mycobacterium tuberculosis*. J Struct Funct Genomics 6:113–118

15. Waldo GS, Standish BM, Berendzen J, Terwilliger TC (1999) Rapid protein folding assay using green fluorescent protein. Nat Biotechnol 17:691–695

16. Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24:79–88