Center for Eukaryotic Structural Genomics

Protein Structure Initiative

A SEMI-AUTOMATED PROTOCOL FOR THE PRODUCTION OF PROTEIN SAMPLES SUITABLE FOR NMR STRUCTURE DETERMINATION

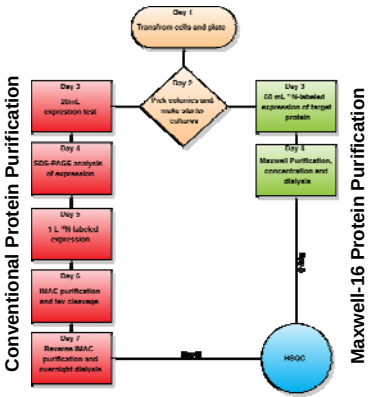
Davin R. Jensen, Christopher Woytovich, Margie Li, Francis C. Peterson, and Brian F. Volkman

Medical College of Wisconsin, Department of Biochemistry, 8701 Watertown Plank Road, Milwaukee, WI USA 53226, http://www.uwstructuralgenomics.org

Abstract

Structural analysis of uniformly folded proteins requires a substantial investment of time, labor and expense. Therefore, before determining the structure of a protein by NMR, a process of protein production, purification, and fold-stability assessment to verify the protein's candidacy for these studies typically ensues. Despite the reliability of conventional processes for determining the suitability of a protein for structural studies, the process is costly, time consuming, and susceptible to the loss of protein at each step. Here we present a semi-automated protocol for the production and purification of proteins using the Maxwell-16, a commercially available benchtop robot. This protocol allows for the efficient production, purification, and NMR screening of the protein samples necessary to evaluate the fold state of the protein. To validate our protocol, we selected a set of eight control proteins and 18 new proteins to put through our semi-automated protocol. In the span of one week, we were able to express, purify, and screen 30 different ^{15}N-labeled proteins. Protein yields from a single channel of the Maxwell-16 were sufficient to prepare 200 µL protein samples for 2D ^{1}H-^{15}N-HSQC NMR using a 3 mm NMR tube. The results for the control group of proteins agreed with our previous observations when using a conventional process for the production and purification of proteins. Analysis of the new protein targets identified three potential protein targets that included the second PDZ domain (PDZ2) of human Par-3 (hPar3). To evaluate the utility of this semi-automated protocol for the production of NMR samples necessary for structure determination, we selected PDZ2 of hPar-3 and prepared a ^{15}N,^{13}C-labeled sample for structural analysis using the Maxwell-16. The resulting protein sample (200 µL volume in a 3 mm NMR tube) was sufficient to collect all the NMR data necessary to determine the structure of PDZ2 from hPar-3. The application of this semi-automated protein production protocol reduces the time and labor typically associated with the preparation of protein samples.

Introduction



Conventional protocols for the production of ^{15}N-labled proteins for 2D-NMR screening requires a substantial investment of resources and is typically done in two stages. Initially, a small-scale expression study (50 mL of culture) is conducted and analyzed by SDS-PAGE to evaluate a protein's expression level and solubility. If these studies indicate that a protein is suitable for further study, then a 1 L production of ^{15}N-labeled protein is initiated and purified by metal affinity chromatography in the second stage. The purified protein sample is then evaluated by 2D NMR to determine its suitability for NMR structure determination. Altogether, this process can span two weeks and is not easily parallelized as a result of the large volumes involved in the production of ^{15}N-labled proteins.

The Maxwell-16 (Promega, Madison, WI) is a small benchtop robot that enables the simultaneous lysis and purification of up to sixteen protein samples in less than 1 hour. The automation of protein purification and the parallel processing ability of the Maxwell-16 prompted us to investigate its utility in preparing ^{15}N-labled protein samples for evaluation by 2D NMR. Here we present an optimized protein purification protocol that enables the production of ^{15}N-labeled protein samples suitable for evaluation by 2D NMR from a single channel of the Maxwell-16 using 60 mL of culture. The low culture volumes required to prepare samples using the Maxwell-16 allowed us to combine expression testing and sample preparation in a single step, significantly reducing the time required to screen samples. Using the streamlined process presented here, we are able to screen ~30 protein samples in the span of one week.

Optimization of the Maxwell-16 Protocol

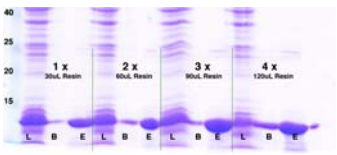


Figure 1. Increasing amounts of MagneHis™ Ni-resin enhances protein yields. Our recreation of the Maxwell-16 Polyhistidine Protein Purification Kit initially used 30 µL of MagneHis™ Ni-resin (1 X lanes). SDS-PAGE analysis of the depleted lysate (L), protein remaining on the beads (B) and the elution (E) fractions indicated that 30 OD_{600} of culture exceeds the capacity of the resin. To determine if we could extract more protein from the lysate, we added increasing amounts of MagneHis™ resin with a constant culture volume. At 120 µL of resin, the depleted lysate retained a small amount of the target protein.

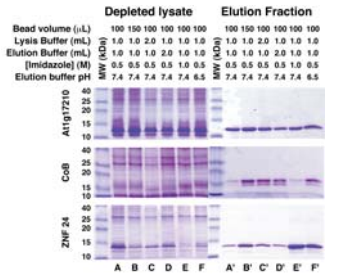


Figure 2. Optimization of Buffer conditions. Protein yields were further enhanced by optimizing the volume of lysis buffer, beads, and elution buffer used. Additionally, we investigated how the elution buffer pH and imidazole concentration influenced release of protein from the MagneHis™ Ni-resin. SDS-PAGE analysis of the three protein targets used in the study reveal target to target variations that would be in a large group of proteins. However, the results indicate that using 150 µL of MagneHis™ Ni-resin and lowering the elution buffer pH to 6.5 yields the best overall results. All studies were conducted with culture volumes corresponding to 60 OD_{600}.

Validation of the Optimized Protocol

Construct	Temp	Expression	Purification	HSQC
CoB	37 °C	+++	+++	+
CoB	15 °C	+++	+++	+
At3g17210	37 °C	+++	+++	+
At3g17210	15 °C	+++	+++	+
Znf24 8HT	37 °C	++	++	-
Znf24 8HT	15 °C	++	++	-
At5g22580	37 °C	+++	+++	+
At5g22580	15 °C	+++	+++	+
At3g29075	37 °C	++	++	unfolded
At3g29075	15 °C	+	+	unfolded
At3g05570	37 °C	++	+	unfolded
At1g16640 8HT	37 °C	+	+	-
At1g16640 8HT	15 °C	+	+	-
At1g16640 GB1	37 °C	+++	+++	-
At1g16640 GB1	15 °C	+++	+++	-
At1g20490	37 °C	+	-	-
At1g20490	15 °C	+	-	-

A control workgroup of 8 proteins was used to validate the optimized protocol and included three unfolded proteins (At1g20940, At3g29075, and At3g05570) and five proteins for which we had previously determined structures. Proteins were expressed with ^{15}N isotopic labeling and purified using the optimized Maxwell-16 protocol from a culture volume corresponding to 60 OD_{600}. Purified proteins were buffer exchanged into a standard NMR buffer (20 mM sodium phosphate, pH 6.5, containing 50 mM sodium chloride), concentrated to a final volume of ~ 200 µL, and evaluated by 2D NMR (**Figure 3**). Results from the control work group agreed well with the previous results, with the exception of Znf24 and At1g16640. The inconsistencies observed for Znf24 and Atg16640 are potentially a consequence of not removing the purification tags prior to sample preparation.

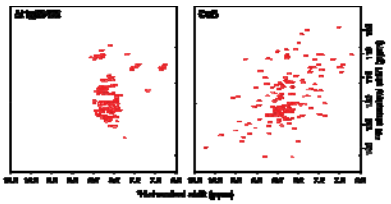


Figure 3. Two-dimensional NMR analysis of control work group proteins. Sample from the control work group were subject to ^{1}H-^{15}N HSQC NMR. Spectra were acquired using a Bruker Avance (Billerica, MA) 600 MHz NMR equipped with a 5 mm TCI Cryoprobe. Data collection was done at 25 °C in 3 mm NMR tubes with 16 transients per FID.

Application to Protein Domains

Construct	Temp	Expression	Purification	HSQC
Par3B 1-135	37 °C	+++	++	-
Par3B 136-270	37 °C	+	+	-
Par3B 136-367	37 °C	+++	-	-
Par3B 271-367	37 °C	++	-	-
Par3B 271-549	37 °C	-	-	-
Par3B 271-689	37 °C	++	-	-
Par3B 451-549	37 °C	+++	+++	+
Par3B 451-689	37 °C	+	+	+/-
Par3B 584-689	37 °C	+	+	-
Par3B 690-840	37 °C	many bands	many bands	unfolded
Par3B 739-958	37 °C	-	-	-
Par3B 942-1025	37 °C	++	+	unfolded
Par3B 1026-1230	37 °C	+	-	-
Par3B 1232-1353	37 °C	+	many bands	unfolded
DH-8HT	37 °C	+++	+	-
DH-8HT	15 °C	++	++	-
DH GB1	37 °C	++	+	-
DH GB1	15 °C	++	++	-
SH3-8HT	37 °C	+++	+	-
SH3-8HT	15 °C	++	++	+
SH3-GB1	37 °C	++	++	-
SH3-GB1	15 °C	++	++	-

The ability to rapidly express and purify multiple protein samples using the Maxwell-16 affords the user the opportunity to optimize their protein expression constructs. This is most helpful when extracting domains from large proteins and the optimal boundaries are not known. To evaluate the utility of the Maxwell-16 in this application, we designed a domain workgroup consisting of a series of constructs that extracted individual or multiple PDZ domains from human Par3B and the Dbl homology domain or Src homology 3 (SH3) domain from human β-PIX. Once the protein expression constructs were in hand, ^{15}N-labeled proteins were expressed and purified using the optimized Maxwell-16 protocol. Of the 18 constructs tested, three failed at the protein expression level, five additional proteins failed at the purification level, and 10 were passed to NMR. Two of the 10 targets were folded by NMR, one was unfolded, and seven yielded no detectable signal.

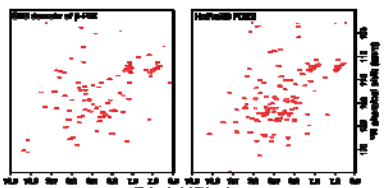


Figure 4. Positive ^{1}H-^{15}N HSQCs from the domain workgroup. ^{1}H-^{15}N HSQCs for the SH3 domain of β–PIX and PDZ2 (residues 451-549) of Par3B. Spectra were acquired using a Bruker Avance 600 MHz NMR equipped with a 5 mm TCI Cryoprobe. Data collection was done at 25 °C in 3 mm NMR tubes with 16 and 64 transients per FID for β–PIX and Par3b, respectively.

Production of ^{15}N/^{13}C Labeled Protein Samples for NMR Structure Determination

To evaluate the utility of the Maxwel-16 in the production ^{15}N/^{13}C-labeled proteins for structure determination, we selected the PDZ2 domain (residues 451-549) of human Par3B as a test case. A one liter culture of ^{15}N/^{13}C Par3B-PDZ2 was expressed and purified using the optimized Maxwell-16 protocol. Briefly, the entire one liter culture was equally divided between the 16 channels and purified. The purified protein was pooled and dialyzed in an appropriate NMR buffer and concentrated to a final concentration of 1 mM in a total volume of 700 µL. The three-dimensional NMR data required for a structure determination was collected on a Bruker Avance 500 MHz NMR at 25 °C using a 3 mm NMR tube (200 µL sample volume). The sample volume used for data collection was equivalent to protein purified from only five channels of the Maxwell-16.

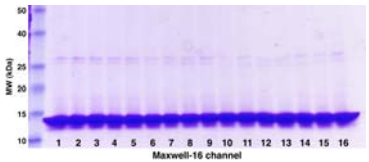


Figure 5. SDS-PAGE analysis of the Par3B-PDZ2 purified using the Maxwell-16 shows that each of the 16 channels performs equally well. Approximately 0.5 mg of purified Par3B-PDZ2 was obtained from each Maxwell-16 channel.

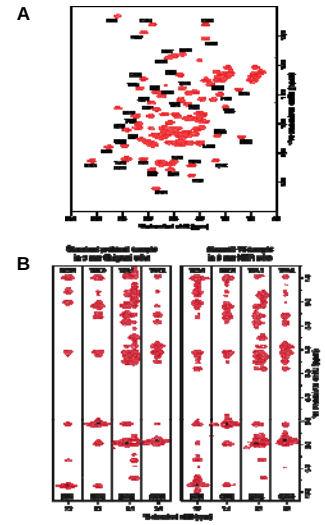


Figure 6. NMR analysis of the Maxwell-16 purified Par3B-PDZ2. (**A**) An assigned ^{1}H-^{15}N-HSQC of Par3B-PDZ2 containing 113 of 114 expected resonances. (**B**) Comparison of four ^{15}N-edited NOESY strips from spectra collected using protein prepared by standard methods or using the Maxwell-16. Data for the standard sample was acquired at 25 °C in a 5 mm Shigemi tube at 600 MHz, while data collection for the Maxwell-16 sample was done at 25 °C in a 3 mm NMR tube at 500 MHz with equivalent data collection times. Comparison of the data sets shows that the reduction in the active NMR sample volume had no detectable effect on the quality of the spectra. Similar results were observed for the standard backbone experiments (data not shown). These results suggest that the ^{15}N/^{13}C-labeled proteins suitable for structure determination could be prepared from as few as 4-5 Maxwell-16 channels.

Conclusion and Future Directions

❖ CONCLUSION: Automated Maxwell-16 purification can support high-throughput NMR-based screening of target proteins for structural genomics.
❖ FUTURE DIRECTION: Modification of Maxwell-16 protocol to enable automated TEV protease digestion and separation of target protein from fusion partners.
❖ FUTURE DIRECTION: Enhancement of NMR sensitivity and additional sample miniaturization through the use of 3 mm or 1.7 mm CryoProbes.

CESG is supported by the National Institute of General Medical Sciences through the Protein Structure Initiative
NIH/NIGMS grant numbers U54 GM074901 (JLM; 07/01/05-06/30/10) and P50 GM064598 (JLM; 01/01/02-08/31/05).