Supporting information

A1

MGSSHHHHHHSSGRENLYFQGLVLVPASQAGVQDQRARFREIFCSILDTRGSDLPDSRPCEEALTRVGSEPGG
TGKPVDLGQSKRRLIGAVVAGVGYECFEQWLQAPGSVAIHLRQFGYDLVQIKVDALSSSANNARQIRDAIMAMP
AEAGPPRLVLIGYSKGAPDILEAMVAYPQIRGRVAAVISAAGSVGGSALANDAEQYQADLLRYFPGATCDSGDGG
AINSLRPATRKSWLAQNRLPSELNYYSLVTFPQPDHISFALKSSYNKLARIDARNDSQVIFYDQVVPGSTLMGYINA
DHWALAVPIARTHSTIASLLVTQNAYPREALAEAMLRFVEEDLAASSR

A2

MGSSHHHHHHSSGRENLYFQGLVPASMTGVVDGRARFREIYCAVTNERGHTLPDYRPCEEALTRLANEPPPTKA
PIDLGVSNSPMRVLIVFGIGAKCIENFIDAQMTVVNHLARFGHKVGSIEVEAISSSTRNASLIRDAVMAMPAATESE
HLVLLGYSKGTPDILEAVTTYPDLQRRVTAVVSVSGSVGGSPLANAATQSMLNLLQYFPNAECDPGDEGGLESLR
PDVRKRWLASHTLPESIRYYSIVTYPDSDHISSILKGSYKKLSQVDSRNDSQMIYYDQLIPGSVLLGYLNADHWAV
AVPFNRSHPFISSTFLGKNAFPREVLIEAIVRYVEEDLNNIQQIRSESK

A4

MGSSHHHHHHSSGRENLYFQGLMLAPASMVGVVDGRARFREIYCAITNERGREMPDYRPCEEALVRLQNEPPP
TGAPVDLGASRSPLRIMIVFGVGAKCIENFIDFQMTVVDHLARFGYKVGILQVEALSSSARNADIIRETVMKMPDQ
NDGMRLVLVGYSKGTPDILEAVATYPDLQQRVSAVVSVAGAVGGTPLANAATQSMLNLLQYFPDADCEPGDEG
ALESLKPEVRKSWLASHSLPDSIRYYTVITYPDAEHLSAILRLSYDKLSQVDSRNDSQVIFYDQVIPSSVLLAYLNAD
HWAVAVPFNRNHPFIASTFVDKNAFPREVLAEAIIRFIEEDLKGR

6-3

MGSSHHHHHHSSGRENLYFQGMILVPASMTGVVDGRARFREIYCAVTNERGHTLPDYRPCEEALTRLANEPPPT
KAPIDLGVSNSPMRVLIVFGIGAKCIENFIDAQMTVVNHLARFGHKVGSIEVEAISSSTRNASLIRDAVMAMPAATE
SEHLVLLGYSKGTPDILEAVTTYPDLQRRVTAVVSVSGSVGGSPLANAATQSMLNLLQYFPNAECDPGDEGGLES
LRPDVRKRWLASHTLPESIRYYSIVTYPDSDHISSILKGSYKKLSQVDSRNDSQMIYYDQLIPGSVLLGYLNADHWA
VAVPFNRSHPFISSTFLGKNAFPREVLIEAIVRYVEEDLNNIQQIRSESK

6-5

MGSSHHHHHHSSGRENLYFQGLMLAPASMVGVVDGRARFREIYCAITNERGREMPDYRPCEEALVRLQNEPPP
TGAPVDLGASRSPLRIMIVFGVGAKCIENFIDFQMTVVDHLARFGYKVGILQVEALSSSARNADIIRETVMTMPDQ
NDGMRLVLVGYSKGTPDILEAVATYPDLQQRVSAVVSVAGAVGGTPLANAATQSMLNLLQYFPDADCEPGDEG
ALESLKPEVRKSWLASHSLPDSIRYYTVITYPDAEHLSAILRLSYDKLSQVDSRNDSQVIFYDQVIPSSVLLAYLNAD
HWAVAVPFNRNHPFIASTFVDKNAFPREVLAEAIIRFIEEDLKGR

6-6

MGSSHHHHHHSSGRENLYFQGLILSPACAAGVIDGRSRFREIFCAVTSERGHKMPDYRPCEDALVRLETESPPT
GVPVDLGVSKSPLRILVVLGIGAKCVENFLDFQTTAAEHLIRFGHKVGILEVEALSSCRRNAGIIREAVMGLPDSEA
GMRMVLLGYSKGTPDILEAVTTYPELQQRISAVVSVAGAVGGSPMANMVEQSTLNLLQYFPGAECDPGDEGGLE
SLRPEVRRRWLAKHTLPKSIRYYSIITYPDADHISSILKSSYKRLSQIDSRNDSQLIFYDQLIPGSVLLGYLNADHWAV
AVPFNRDHPVIASQFVDKNAFPREVLAEAIIRYVEEDLGWR

A2119N

MGSSHHHHHHSSGRENLYFQG**SAPKPMIQYSTDTPP**LVPASMTGVVDGRARFREIYCAVTNERGHTLPDYRPC
EEALTRLANEPPPTKAPIDLGVSNSPMRVLIVFGIGAKCIENFIDAQMTVVNHLARFGHKVGSIEVEAISSSTRNASLI
RDAVMAMPAATESEHLVLLGYSKGTPDILEAVTTYPDLQRRVTAVVSVSGSVGGSPLANAATQSMLNLLQYFPNA
ECDPGDEGGLESLRPDVRKRWLASHTLPESIRYYSIVTYPDSDHISSILKGSYKKLSQVDSRNDSQMIYYDQLIPGS
VLLGYLNADHWAVAVPFNRSHPFISSTFLGKNAFPREVLIEAIVRYVEEDLNNIQQIRSESK

A2WS

MGSSHHHHHHSSGRENLYFQG**SAPKALIPHSLDTPPMI**LVPASMTGVVDGRARFREIYCAVTNERGHTLPDYR
PCEEALTRLANEPPPTKAPIDLGVSNSPMRVLIVFGIGAKCIENFIDAQMTVVNHLARFGHKVGSIEVEAISSSTRNA
SLIRDAVMAMPAATESEHLVLLGYSKGTPDILEAVTTYPDLQRRVTAVVSVSGSVGGSPLANAATQSMLNLLQYFP
NAECDPGDEGGLESLRPDVRKRWLASHTLPESIRYYSIVTYPDSDHISSILKGSYKKLSQVDSRNDSQMIYYDQLIP
GSVLLGYLNADHWAVAVPFNRSHPFISSTFLGKNAFPREVLIEAIVRYVEEDLNNIQQIRSESK

Figure S1 Primary sequences of the selected enzymes (**A1**, **A2**, **A4**, **6-3**, **6-5**, **6-6**, **A2WS**, and **A2119N**), each containing an N-terminal His_6-tag followed by a TEV protease cleavage site (highlighted in yellow). In **A2119N**, a 15-amino acid N-terminal extension (SAPKPMIQYSTDTPP), derived from a previously engineered enzyme (unpublished results), is shown in bold cyan. In **A2WS**, the wild-type N-terminal sequence is shown in bold cyan. All sequences were codon-optimized for *E. coli* BL21(DE3) expression and synthesized by Twist Bioscience. The genes were cloned into pET-28a(+) vectors, providing an N-terminal His_6-tag for purification.

Table S1. Calibration data for size-exclusion chromatography (SEC) using molecular weight standards on a Superdex 200 10/300 GL column.

Protein	M_r (theoretical)/ 10^3	M_r (experimental)/ 10^3	log (M_r / 10^3)	V_e / mL	K_{av}
Tireoglobulin	670	841	2.826	8.79	0.013358269
γ-globulin	158	144	2.199	12.41	0.255142933
ovalbumin	44	50	1.643	14.57	0.399412236
mioglobin	17	10	1.230	17.92	0.623163238
vitamin B12	1.35	1.8	0.130	21.36	0.852925461
A2	38.7	27	1.431	15.83	0.483569329
A2WS	40.5	52	1.713	14.5	0.394736842
6-5	38.2	24	1.376	16.09	0.500935079

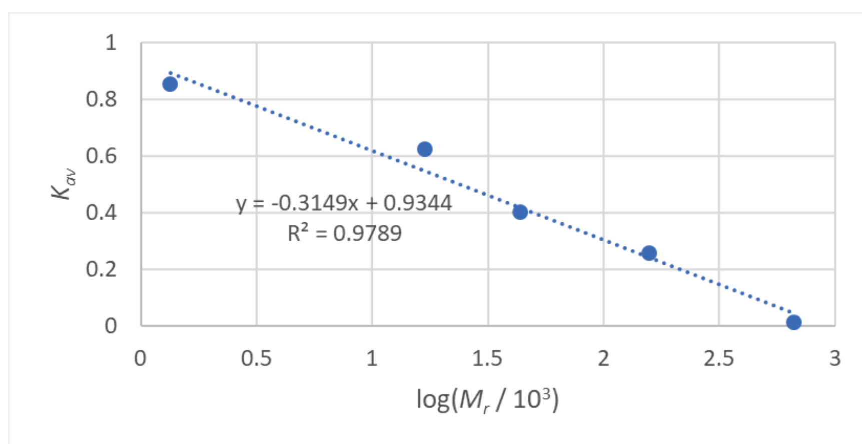

Figure S2. Plot of K_{av} versus log(M_r / 10^3)

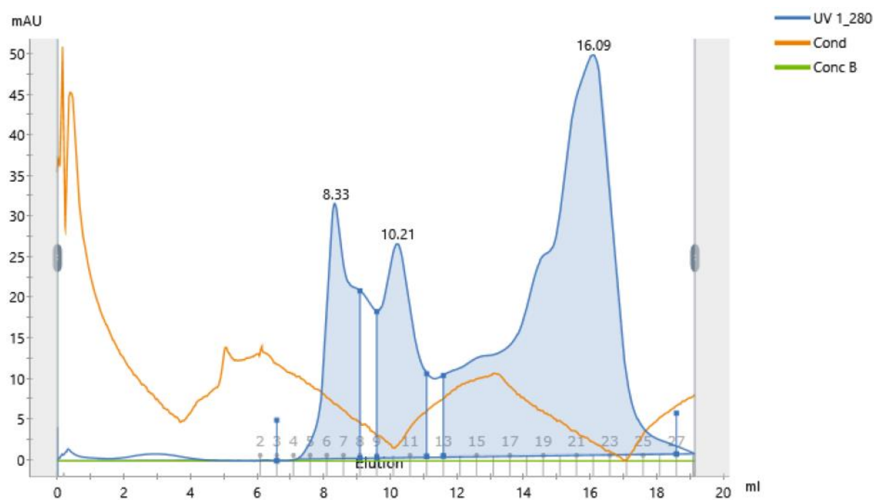

Figure S3. Size-exclusion chromatography (SEC) profile of enzyme **6-5**. Collected fractions are labeled 2–27 (shown in grey); fractions 20–23 correspond to the monomeric form.

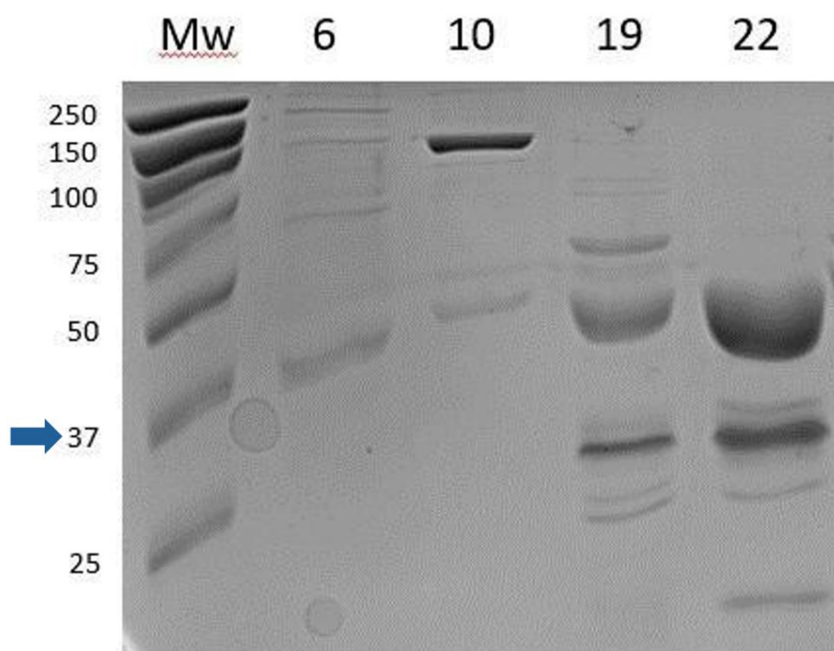

Figure S4. SDS-PAGE analysis of **6-5** protein fractions after size-exclusion chromatography (M_w- molecular weight marker). The blue arrow indicates the expected band of the overexpressed protein. The predicted molecular mass of the 6-5 is 38.2 kDa. Protein bands were visualized using a CCD imaging system.

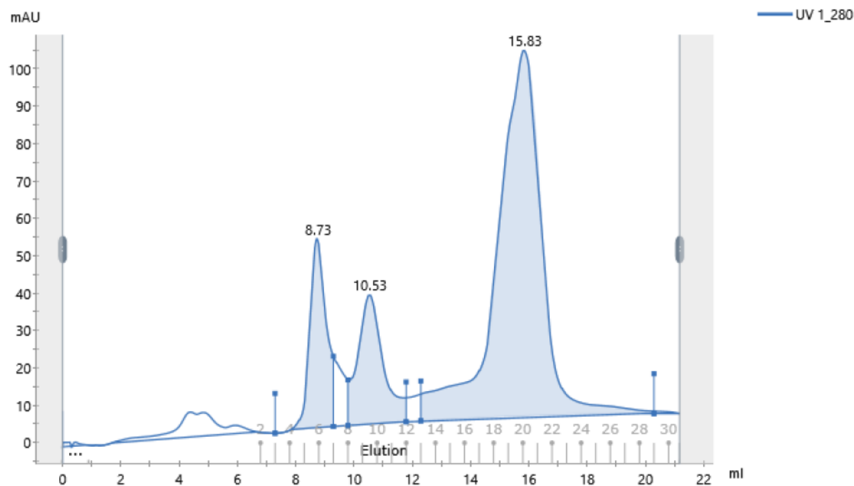

Figure S5. Size-exclusion chromatography (SEC) profile of enzyme **A2**. Collected fractions are labeled 2–30 (shown in grey); fractions 17–22 correspond to the monomeric form.

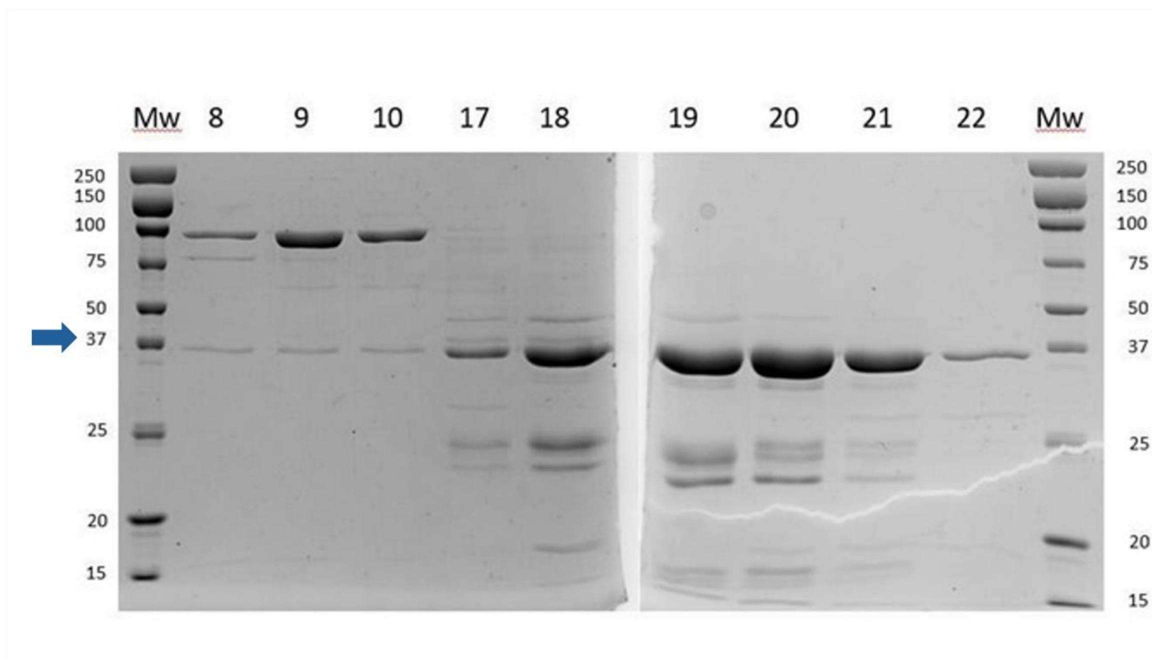

Figure S6. SDS-PAGE analysis of **A2** protein fractions after size-exclusion chromatography (M_w- molecular weight marker). The blue arrow indicates the expected band of the overexpressed protein. The predicted molecular mass of the **A2** is 38.7 kDa. Protein bands were visualized using a CCD imaging system.

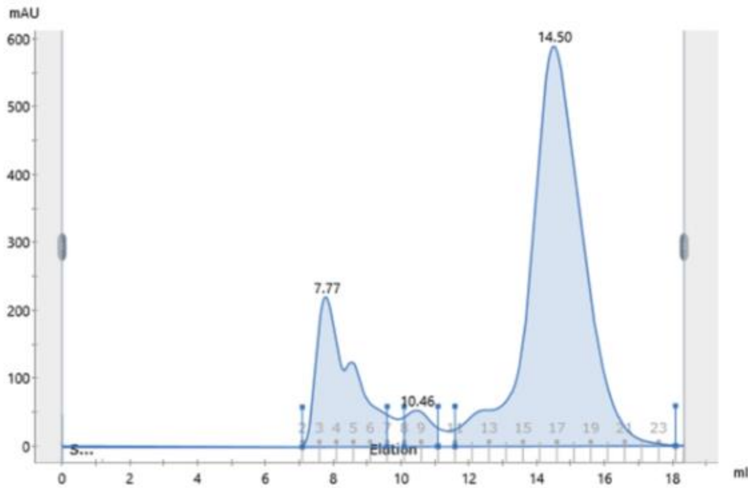

Figure S7. Size-exclusion chromatography (SEC) profile of enzyme **A2WS**. Collected fractions are labeled 2–23 (shown in grey); fractions 14–21 correspond to the monomeric form.

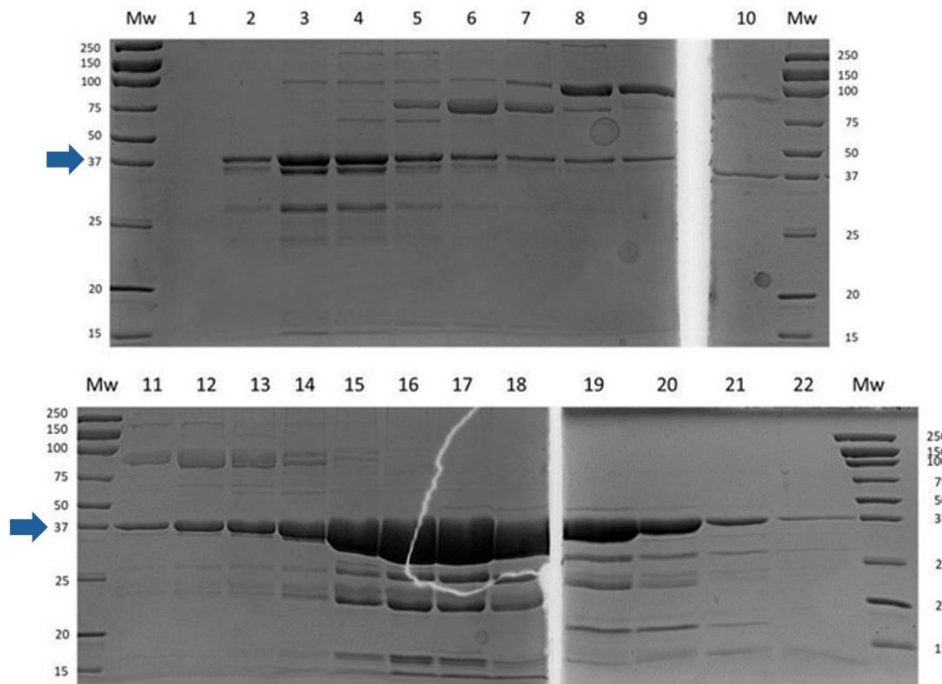

Figure S8. SDS-PAGE analysis of **A2WS** protein fractions after size-exclusion chromatography (M_w- molecular weight marker). The blue arrows indicate the expected band of the overexpressed protein. The predicted molecular mass of the **A2WS** is 40.5 kDa. Protein bands were visualized using a CCD imaging system.

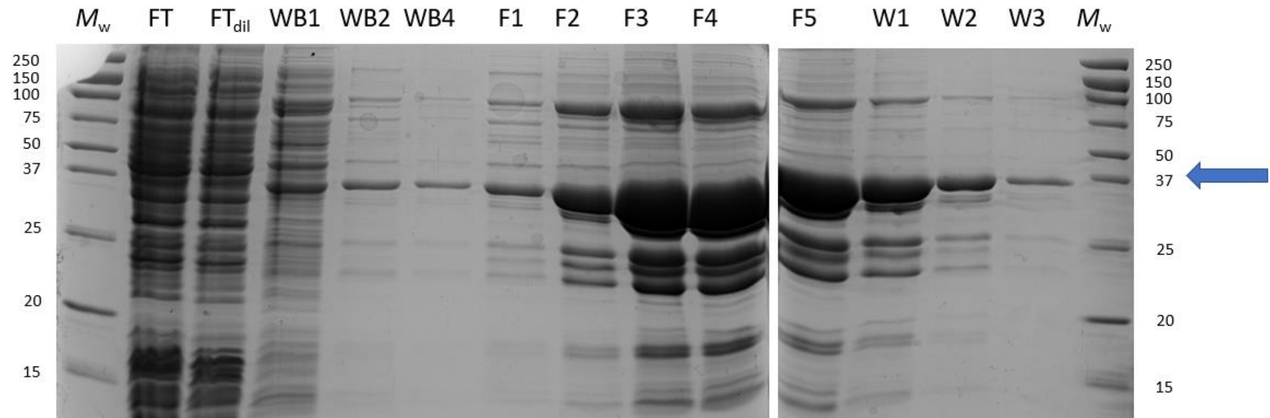

Figure S9. SDS-PAGE analysis of fractions after Ni-NTA affinity chromatography of the soluble protein extract of *E. coli* containing the plasmid for the **A2** protein from the 0.5 L culture induced with IPTG at 15 °C. M_w – molecular weight marker, FT – flow-through fraction from the Ni^{2+}-NTA column, FT_{dil} – diluted flow-through fraction from the Ni^{2+}-NTA column, WB 1–4 – fractions obtained by elution with buffer B from the column (imidazole concentration = 40 mmol/dm³), F 1–5 – fractions obtained by elution with buffer C (imidazole concentration = 200 mmol/dm³), W 1–3 – fractions obtained by elution with buffer 500 (imidazole concentration = 500 mmol/dm³). The theoretical molecular mass of **A2** protein is 38.7 kDa and is indicated with a blue arrow. Gels were visualized using a CCD camera.

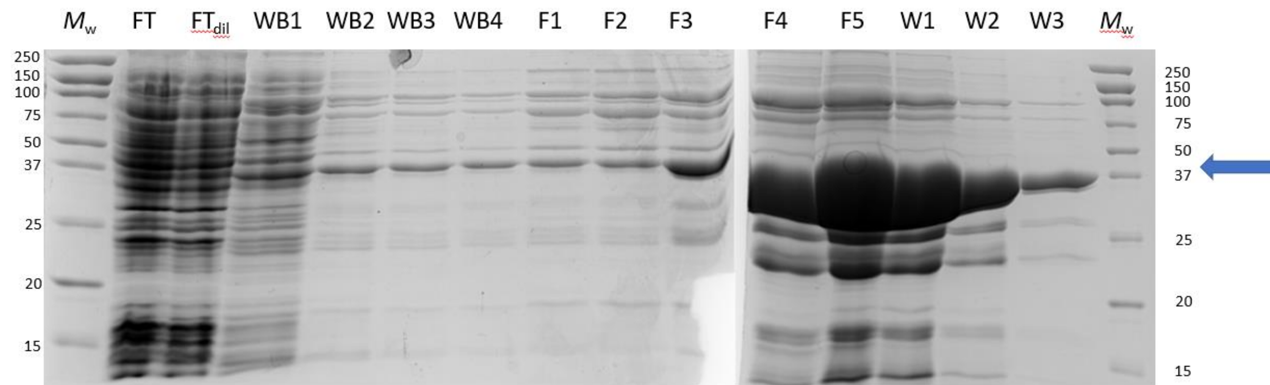

Figure S10. SDS-PAGE analysis of fractions after Ni-NTA affinity chromatography of the soluble protein extract of *E. coli* containing the plasmid for the **A2WS** protein from the 6 L culture induced with IPTG at 15 °C. M_w – molecular weight marker, FT – flow-through fraction from the Ni^{2+}-NTA column, FT_{dil} – diluted flow-through fraction from the Ni^{2+}-NTA column, WB 1–4 – fractions obtained by elution with buffer B from the column (imidazole concentration = 40 mmol/dm³), F 1–5 – fractions obtained by elution with buffer C (imidazole concentration = 200 mmol/dm³), W 1–3 – fractions obtained by elution with buffer 500 (imidazole

concentration = 500 mmol/dm³). The theoretical molecular mass of **A2WS** protein is 40.5 kDa and is indicated with a blue arrow. Gels were visualized using a CCD camera.

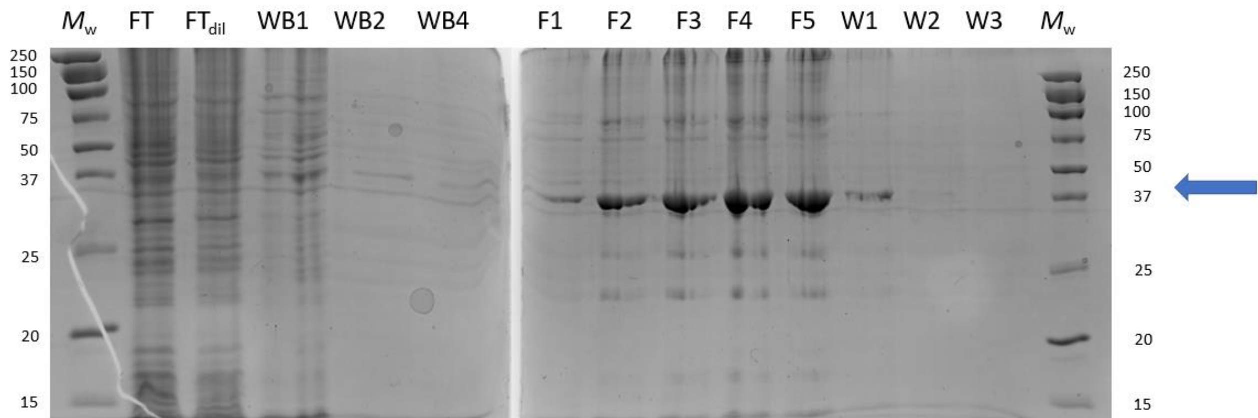

Figure S11. SDS-PAGE analysis of fractions after Ni-NTA affinity chromatography of the soluble protein extract of *E. coli* containing the plasmid for the **A2119N** protein from the 1 L culture induced with IPTG at 15 °C. M_w – molecular weight marker, FT – flow-through fraction from the Ni^{2+}-NTA column, FT_{dil} – diluted flow-through fraction from the Ni^{2+}-NTA column, WB 1–4 – fractions obtained by elution with buffer B from the column (imidazole concentration = 40 mmol/dm³), F 1–5 – fractions obtained by elution with buffer C (imidazole concentration = 200 mmol/dm³), W 1–3 – fractions obtained by elution with buffer 500 (imidazole concentration = 500 mmol/dm³). The theoretical molecular mass of **A2119N** protein is 40.3 kDa and is indicated with a blue arrow. Gels were visualized using a CCD camera.

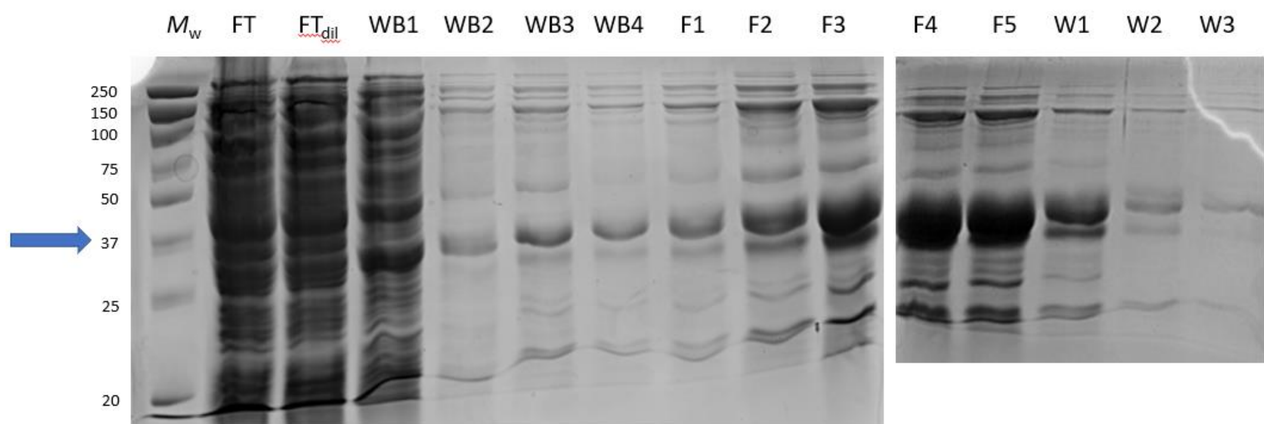

Figure S12. SDS-PAGE analysis of fractions after Ni-NTA affinity chromatography of the soluble protein extract of *E. coli* containing the plasmid for the **6-5** protein from the 1 L culture induced with IPTG at 15 °C. M_w – molecular weight marker, FT – flow-through fraction from the Ni^{2+}-NTA column, FT_{dil} – diluted flow-through

fraction from the Ni^{2+}-NTA column, WB 1–4 – fractions obtained by elution with buffer B from the column (imidazole concentration = 40 mmol/dm^3), F 1–5 – fractions obtained by elution with buffer C (imidazole concentration = 200 mmol/dm^3), W 1–3 – fractions obtained by elution with buffer 500 (imidazole concentration = 500 mmol/dm^3). The theoretical molecular mass of **6-5** protein is 38.2 kDa and is indicated with a blue arrow. Gels were visualized using a CCD camera.