

Prokaryotic Expression and Purification of the hSox2-HMG Domain

Lijie Yang^{1, 2, #}, Wansen Tan^{1, #} and Jingjun Hong^{1, 2, #, *}

¹Institute of Health Sciences and Technology (IHST), Institutes of Physical Sciences and Information Technology, Anhui University, Hefei, China

²School of Life Sciences and Medical Engineering, Anhui University, Hefei, China

*For correspondence: jjhong@ahu.edu.cn

#Contributed equally to this work

Abstract

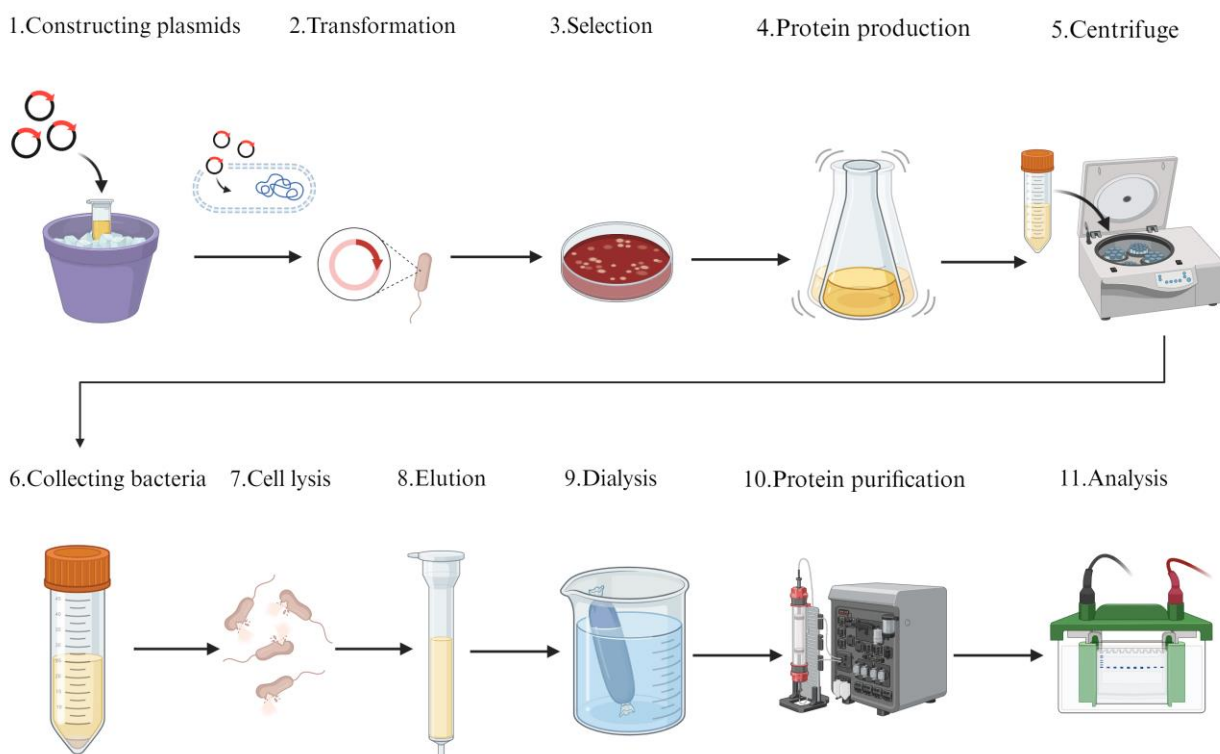
The Sox (SRY-related HMG-box) protein family plays a crucial role in cellular differentiation, development, and gene regulation, with the HMG (high-mobility group) domain responsible for DNA binding and transcriptional regulation. Proteins in the SOX gene family contain an HMG domain that shares 50% homology with the HMG domain of the sex-determining factor SRY gene. The SOX gene family comprises 30 proteins, which are classified into 10 groups (A–H). As a member of this family, hSox2 has been shown to be involved in various biological processes, but its specific function remains unclear. Previous studies have used eukaryotic expression systems, GST-tag purification, and bacterial inclusion body refolding techniques to produce Sox family proteins. However, these methods are often limited by issues such as low yield, incorrect folding, or inefficient purification, restricting their application in functional and structural studies. In this study, a prokaryotic expression system for the hSox2-HMG domain was constructed using the pET22b vector and *Escherichia coli* BL21(DE3) as the host strain. Protein expression was induced by IPTG, and initial purification was performed using Ni-NTA affinity chromatography, followed by ultrafiltration concentration and size exclusion chromatography to improve purity. By optimizing lysis and elution conditions, we successfully obtained hSox2-HMG protein with high expression levels and purity. This method provides a cost-effective and scalable strategy for hSox2-HMG production, ensuring high purity and correct folding of the protein. The optimized experimental protocol lays a foundation for structural and functional studies of hSox2-HMG.

Key features

- The hSox2-HMG protein was expressed in *Escherichia coli* BL21(DE3) using the pET22b vector and IPTG induction, resulting in high-yield recombinant protein.
- Ni-NTA affinity chromatography was employed for protein purification and combined with ultrafiltration concentration and size exclusion chromatography to enhance purity and ensure correct folding.
- The established workflow provides an efficient, cost-effective, and scalable strategy for hSox2-HMG production, suitable for structural and functional studies.
- The purified hSox2-HMG protein maintains structural integrity and can be used for further investigation of DNA-binding properties and regulatory functions.

Keywords: hSox2-HMG, Prokaryotic expression, Protein purification, Ni-NTA affinity chromatography, IPTG induction, Size exclusion chromatography, Recombinant protein, Functional studies

Graphical overview



Flowchart of prokaryotic expression and purification of the hSOX2-HMG domain. Constructing plasmids: synthesized by General Biol (Anhui) Co. Transformation: *E. coli* BL21(DE3) cells and the expression plasmid (hSOX2-HMG-pET22b) were extracted on ice. Selection: all colonies on the LB plate were scraped for protein expression. Protein production: protein expression was performed using shake-flask culture. Centrifuge: cells were separated from the culture medium by centrifugation. Collecting bacteria: discard the culture medium. Cell lysis: cells were disrupted by sonication with an ultrasonic cell disruptor. Elution: through Ni-NTA. Dialysis: replacement buffer. Protein purification: through size exclusion. Analysis: through SDS-PAGE.

Background

In the regulation of embryogenesis, multiple transcription factors and proteins interacting with DNA play a crucial role, among which is the SOX2 within the SOX B1 subfamily. The SOX gene family is classified into 10 groups (A–H) [1]. The high-mobility group HMG-box (HMG) domain serves as the primary structural basis for DNA binding in SOX family transcription factors and is highly conserved [2]. The pioneer transcription factor OCT4 binds to the octameric motif (5'-ATTTGCAT-3') and forms a trimeric complex with SOX2 or SOX15 on DNA, regulating the expression of genes involved in embryonic development, such as YES1, FGF4, UTF1, and ZFP206. It is crucial for early embryogenesis and the pluripotency of embryonic stem cells (ESCs) [3]. Shinya Yamanaka first demonstrated in 2007 that SOX2, in combination with three other transcription factors (Oct4/POU5F1, Klf4, and c-Myc/MYC), could cooperatively reprogram mouse fibroblast cells into induced pluripotent stem cells (iPSCs) through somatic cell reprogramming in vitro. When combined, these factors are sufficient to reprogram differentiated cells into an embryonic-like state, generating induced pluripotent stem (iPS) cells. iPS cells exhibit the morphology and growth characteristics of ESCs and express ESC marker genes [3]. SOX2 has emerged as the most extensively investigated transcription factor within the SOX B1 subfamily. However, conventional eukaryotic expression systems for the production of SOX2 are typically characterized by low yields, high production costs [4], and a requirement for expensive affinity chromatography media during purification. In contrast, the present study introduces an efficient prokaryotic alternative. Utilizing *Escherichia coli* as the expression host, we achieved predominantly soluble expression of SOX2 HMG, thereby circumventing the necessity for the inclusion of body refolding.

The purification protocol delineated herein consistently affords high-purity SOX2 HMG protein, establishing a robust and cost-effective methodology for its production.

Materials and reagents

Biological materials

Strains

1. *Escherichia coli* BL21 (DE3) competent cells (Sangon Biotech, catalog number: B528414-0100)

Plasmids

The codon-optimized cDNA sequence (see specific sequences below) for hSOX2 HMG domain was synthesized by General Biol (Anhui) Co. and was subsequently cloned into the pET22b (+) expression vector [from General Biol (Anhui) Co., sequence verified] with *Nde* I and *Xho* I cloning sites, to produce hSOX2-HMG-6×His-pET22b expression plasmid.

Codon-optimized cDNA base sequence:

CATATGGATCGTGTGAAACGTCCGATGAATGCATTTCATGGTGTGGAGCCGTGGTCAGCGCCGCAAAATGGC
ACAGGAAAATCCGAAAATGCATAATAGTGAAATCAGTAAACGCCTGGGTGCAGAAATGGAACTGCTGAGC
GAAACCGAAAAACGTCCGTTTATTGATGAAGCAAAACGCCTGCGCGCCCTGCACATGAAAGAACATCCGGA
TTATAAATATCGCCCCGCGTCGCAAAACCAAAACCCTCGAG

Underlined are the hSOX2-HMG DNA base sequences.

SOX2 binding DNA sequence:

P1: 5'-GGAAACAATGGA-3' (transcription factor SOX2 binds to DNA at the 5'-ACAATG-3' consensus sequence)

P2: 5'-TCCATTGTTCC-3'

The DNA base sequence recognized by hSOX2 was synthesized by Sangon Biotech. Double-stranded DNA was prepared through annealing.

Reagents

1. Tryptone (Sangon Biotech, CAS No.: 73049-73-7)
2. Yeast extract (Sangon Biotech, CAS No.: 8013-01-2)
3. NaCl (MACKLIN, CAS No.: 7647-14-5)
4. Agarose (Biowest, CAS No.: 9012-36-6)
5. Sodium phosphate dibasic (Na₂HPO₄) (MACKLIN, CAS No.: 7558-79-4)
6. Potassium phosphate monobasic (KH₂PO₄) (MACKLIN, CAS No.: 7778-77-0)
7. Ampicillin sodium (Amp) (BBI, CAS No.: 69-52-3)
8. Isopropyl-β-D-thiogalactopyranoside (IPTG) (Aladdin, CAS No.: 367-93-1)
9. Ethylenediaminetetraacetic acid (EDTA) (Aladdin, CAS No.: 60-00-4)
10. Ni-NTA agarose (Qiagen, catalog number: 30230)
11. Glycerol (Aladdin, CAS No.: 56-81-5)
12. Phenylmethylsulfonyl fluoride (PMSF) (Aladdin, CAS No.: 329-98-6)
13. Urea (MACKLIN, CAS No.: 57-13-6)
14. Imidazole (MACKLIN, CAS No.: 288-32-4)
15. Sodium dodecyl sulfate (SDS) (BioFroxx, CAS No.: 151-21-3)
16. Ammonium persulfate (APS) (Aladdin, CAS No.: 7727-54-0)
17. Hydroxymethyl aminomethane (Tris) (MACKLIN, CAS No.: 77-86-1)
18. Tetramethylethylenediamine (TEMED) (Aladdin, CAS No.: 110-18-9)
19. Acrylamide (Sangon Biotech, CAS No.: 79-06-1)
20. N,N'-Methylenebisacrylamide (MACKLIN, CAS No.: 110-26-9)
21. Coomassie brilliant blue G-250 (BBI, CAS No.: 6104-58-1)
22. Isopropyl alcohol (Aladdin, CAS No.: 67-63-0)
23. Acetic acid (Sangon Biotech, CAS No.: 64-19-17)

24. β -mercaptoethanol (β -ME) (Aladdin, CAS No.: 60-24-2)
25. NaOH (BBI, CAS No.: 1310-73-2)
26. Boric acid (MACKLIN, CAS No.: 10043-35-3)
27. Glycine (Meilunbio, CAS No.: 56-40-6)
28. Prestained protein ladder (10–180 kDa) (SmartBuffers, catalog number: N6616)
29. Bromophenol blue (Aladdin, CAS No.: 115-39-9)
30. Enhanced BCA Protein Assay kit (Biosharp, catalog number: BL1054C)
31. Glycine (MACKLIN, CAS No.: 50-01-1)
32. HCl (Sinopharm Chemical Reagent Co., Ltd., CAS No.: 7647-01-0)
33. DL-Dithiothreitol (Aladdin, CAS No.: 3483-12-3)
34. Agar (Sangon Biotech, CAS No.: 9002-18-0)

Solutions

1. LB liquid/agar solid medium (see Recipes)
2. Ni affinity chromatography purification binding buffer (washing buffer 1) (see Recipes)
3. Ni affinity chromatography purification washing buffer 2 (see Recipes)
4. Ni affinity chromatography purification washing buffer 3 (see Recipes)
5. Ni affinity chromatography purification elution buffer 1 (see Recipes)
6. Ni affinity chromatography purification elution buffer 2 (see Recipes)
7. 6% PAGE (see Recipes)
8. 15% SDS-PAGE (see Recipes)
9. 50% (v/v) glycerol (see Recipes)
10. Ampicillin stock (see Recipes)
11. 0.5 M IPTG (see Recipes)
12. Gel filtration buffer (see Recipes)
13. Coomassie brilliant blue stain (see Recipes)
14. 5 \times TBE (see Recipes)
15. 0.5 M EDTA pH 8.0 (see Recipes)
16. 5 \times Running buffer (see Recipes)
17. 5 \times SDS loading buffer (see Recipes)
18. 6 M Guanidine (see Recipes)
19. 5 M imidazole pH 7.2 or 8.0 (see Recipes)
20. 30% Acrylamide (see Recipes)
21. 1.5 M Tris-HCl pH 8.8 (see Recipes)
22. 1.0 M Tris-HCl pH 6.8 (see Recipes)
23. 10% SDS (see Recipes)
24. 10% APS (see Recipes)
25. 1 M DL-Dithiothreitol (DTT) (see Recipes)

Recipes

1. LB liquid/agar solid medium (100 mL)

Reagent	Final concentration	Quantity or Volume
Tryptone	7 g/L	0.7 g
Yeast extract	3.5 g/L	0.35 g
Sodium chloride	10 g/L	1 g
Agar powder (alternative)	20 g/L	2 g
H ₂ O	n/a	100 mL

Prepare fresh and use immediately. Make up the volume to 1,00 mL. After mixing all reagents thoroughly, sterilize by autoclaving on liquid setting (121 °C for 20 min).

2. Ni affinity chromatography purification binding buffer (washing buffer 1) (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole pH 8.0	10 mM	2 mL
NaCl	0.5 M	125 mL
Tris-HCl pH 8.0	20 mM	20 mL
Glycerol	0.2%	4 mL
β-ME	1 mM	1 mL
Urea	0.1 M	14 mL
H ₂ O	n/a	To 1,000 mL

Use a preconfigured solution: 5 M imidazole pH 8.0, 4 M NaCl, 1 M Tris-HCl pH 8.0, 50% glycerol, 1 M β-ME, and 7 M Urea. Add ultrapure water to a final volume of 1,000 mL. Store at 4 °C for up to four weeks.

3. Ni affinity chromatography purification washing buffer 2 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole pH 8.0	25 mM	5 mL
NaCl	0.5 M	125 mL
Tris-HCl pH 8.0	20 mM	20 mL
Glycerol	0.2%	4 mL
β-ME	1 mM	1 mL
Urea	0.1 M	14 mL
H ₂ O	n/a	To 1,000 mL

Use a preconfigured solution: 5 M imidazole pH 8.0, 4 M NaCl, 1 M Tris-HCl pH 8.0, 50% glycerol, 1 M β-ME, and 7 M Urea. Add ultrapure water to a final volume of 1,000 mL. Store at 4 °C for up to four weeks.

4. Ni affinity chromatography purification washing buffer 3 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole pH 8.0	50 mM	10 mL
NaCl	0.5 M	125 mL
Tris-HCl pH 8.0	20 mM	20 mL
Glycerol	0.2%	4 mL
β-ME	1 mM	1 mL
Urea	0.1 M	14 mL
H ₂ O	n/a	To 1,000 mL

Use a preconfigured solution: 5 M imidazole pH 8.0, 4 M NaCl, 1 M Tris-HCl pH 8.0, 50% glycerol, 1 M β-ME, and 7 M Urea. Add ultrapure water to a final volume of 1,000 mL. Store at 4 °C for up to four weeks.

5. Ni affinity chromatography purification elution buffer 1 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole pH 7.2	250 mM	50 mL
NaCl	0.5 M	125 mL
Tris-HCl pH 7.2	20 mM	20 mL
Glycerol	0.2%	4 mL
β-ME	1 mM	1 mL
Urea	0.1 M	14 mL
H ₂ O	n/a	To 1,000 mL

Use a preconfigured solution: 5 M imidazole pH 7.2, 4 M NaCl, 1 M Tris-HCl pH 7.2, 50% glycerol, 1 M β-ME, and 7 M Urea. Add ultrapure water to a final volume of 1,000 mL. Store at 4 °C for up to four weeks.

6. Ni affinity chromatography purification elution buffer 2 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole pH 7.2	500 mM	100 mL
NaCl	0.5 M	125 mL
Tris-HCl pH 7.2	20 mM	20 mL
Glycerol	0.2%	4 mL

β -ME	1 mM	1 mL
Urea	0.1 M	14 mL
H ₂ O	n/a	To 1,000 mL

Use a preconfigured solution: 5 M imidazole pH 7.2, 4 M NaCl, 1 M Tris-HCl pH 7.2, 50% glycerol, 1 M β -ME, and 7 M Urea. Add ultrapure water to a final volume of 1,000 mL. Store at 4 °C for up to four weeks.

7. 6% PAGE (16 mL)

Reagent	Volume
5× TBE	1.6 mL
30% Acrylamide	3.2 mL
10% APS	150 μ L
TEMED	20 μ L
H ₂ O	To 16 mL

Prepare fresh and use immediately. Add 5× TBE, 30% Acrylamide, 10% APS, TEMED, and H₂O in that order. Insert a 15-well 1 mm comb and let it solidify at room temperature for about 2 h.

8. 15% SDS-PAGE

Reagent	Volume (15 mL) Separation gel	Reagent	Volume (6 mL) Stacking gel
ddH ₂ O	3.4 mL	ddH ₂ O	4.2 mL
30% Acrylamide	7.5 mL	30% Acrylamide	1 mL
1.5 M Tris-HCl pH 8.8	3.8 mL	1.0 M Tris-HCl pH 6.8	760 μ L
10% SDS	150 μ L	10% SDS	60 μ L
10% APS	150 μ L	10% APS	60 μ L
TEMED	6 μ L	TEMED	3 μ L

Prepare fresh and use immediately. Add ddH₂O, 30% Acrylamide, 1.0 M Tris-HCl pH 6.8 (1.5 M Tris-HCl pH 8.8), 10% SDS, 10% APS, and TEMED in that order. Insert a 15-well 1 mm comb and let it solidify at room temperature for about 1 h.

9. 50% (v/v) glycerol (100 mL)

Reagent	Final concentration	Quantity or Volume
100% (v/v) glycerol	50% (v/v)	50 mL
H ₂ O	n/a	50 mL

Mix 100% (v/v) glycerol and H₂O (pour pure glycerol into the H₂O). Stir for 5 min and then transfer to a 100 mL bottle. Sterilize by autoclaving on liquid setting (121 °C for 20 min). Store at 4 °C for up to four weeks.

10. Ampicillin stock (10 mL)

Reagent	Final concentration	Quantity or Volume
Ampicillin	100 g/L	1 g
H ₂ O	n/a	10 mL

Sterilize by filtering with a 0.2 μ m syringe-end filter. Store in 1 mL aliquots at -20 °C indefinitely.

11. 0.5 M IPTG (10 mL)

Reagent	Final concentration	Quantity or Volume
IPTG	0.5 M	1.19 g
H ₂ O	n/a	10 mL

Sterilize by filtering with a 0.2 μ m syringe-end filter. Store in 1 mL aliquots at -20 °C indefinitely.

12. Gel filtration buffer (1,000 mL)

Reagent	Final concentration	Quantity or Volume
NaCl	0.5 M	125 mL
Tris-HCl pH 7.2	20 mM	20 mL
EDTA	1 mM	2 mL
H ₂ O	n/a	To 1,000 mL

Prepare fresh and use immediately. Use a preconfigured solution: 4 M NaCl, 1 M Tris-HCl pH 7.2, and 0.5 M EDTA. Add ultrapure water to a final volume of 1,000 mL.

13. Coomassie brilliant blue stain

Reagent	Final concentration	Quantity or Volume
Coomassie brilliant blue G-250	0.5 g/L	0.5 g
Isopropyl alcohol	25% (v/v)	250 mL
Acetic acid	10% (v/v)	100 mL
H ₂ O	n/a	To 1,000 mL

Mix 0.5 g of Coomassie brilliant blue G-250, isopropyl alcohol, and acetic acid, and wait until Coomassie brilliant blue G-250 is completely dissolved. Add ultrapure water to a final volume of 1,000 mL. Store at room temperature for up to four weeks.

14. 5× TBE (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Tris	54 g/L	54 g
Boric acid	27.5 g/L	27.5 g
0.5 M EDTA pH 8.0	10 mM	20 mL
H ₂ O	n/a	To 1,000 mL

Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. Store at room temperature for up to four weeks.

15. 0.5 M EDTA pH 8.0 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
EDTA	0.5 M	146.12 g
H ₂ O	n/a	To 1,000 mL

Add solid NaOH and adjust pH to 8.0. Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. Store at 4 °C for up to four weeks.

16. 5× running buffer (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Tris	15 g/L	15 g
SDS	5 g/L	5 g
Glycine	72 g/L	72 g
H ₂ O	n/a	To 1,000 mL

Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. The working concentration was diluted to 1×. Store at room temperature for up to four weeks.

17. 5× SDS loading buffer (100 mL)

Reagent	Final concentration	Quantity or Volume
1 M Tris-HCl pH 6.8	0.25 M	25 mL
10% SDS (w/v)	2% (w/v)	20 mL
50% Glycerol (v/v)	25% (w/v)	50 mL
1 M DTT	20 mM	2 mL
Bromophenol blue	0.2 g/L	0.02 g
H ₂ O	n/a	To 100 mL

Add ultrapure water to a final volume of 100 mL. Store indefinitely at -20 °C.

18. 6 M Guanidine (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Guanidine	6 M	573.18 g
H ₂ O	n/a	To 1000 mL

Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. Store at room temperature for up to four weeks.

19. 5 M imidazole pH 7.2 or 8.0 (500 mL)

Reagent	Final concentration	Quantity or Volume
Imidazole	5 M	170.20 g
H ₂ O	n/a	To 500 mL

Add HCl and adjust pH to 7.2 or 8.0. Add ultrapure water to a final volume of 500 mL. Mix well and filter using a 0.22-μm filter. Store at 4 °C for up to four weeks.

20. 30% Acrylamide (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Acrylamide	292 g/L	292 g
N,N'-Methylenebisacrylamide	8 g/L	8 g
H ₂ O	n/a	To 1,000 mL

Add ultrapure water to a final volume of 1,000 mL. Mix well and store at 4 °C for up to four weeks.

21. 1.5 M Tris-HCl pH 8.8 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Tris	1.5 M	181.71 g
H ₂ O	n/a	To 1,000 mL

Add HCl and adjust pH to 8.8. Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. Store at room temperature for up to four weeks.

22. 1.0 M Tris-HCl pH 6.8 (1,000 mL)

Reagent	Final concentration	Quantity or Volume
Tris	1 M	121.14 g
H ₂ O	n/a	To 1,000 mL

Add HCl and adjust pH to 6.8. Add ultrapure water to a final volume of 1,000 mL. Mix well and filter using a 0.22-μm filter. Store at room temperature for up to four weeks.

23. 10% SDS (1,000 mL)

Reagent	Final concentration	Quantity or Volume
SDS	100 g/L	100 g
H ₂ O	n/a	To 1,000 mL

Add ultrapure water to a final volume of 1,000 mL. Mix well and store at room temperature for up to four weeks.

24. 10% APS (50 mL)

Reagent	Final concentration	Quantity or Volume
APS	100 g/L	5 g
H ₂ O	n/a	To 50 mL

Add ultrapure water to a final volume of 50 mL. Mix well and store at 4 °C for up to four weeks.

25. 1 M DL-Dithiothreitol (DTT) (50 mL)

Reagent	Final concentration	Quantity or Volume
DL-Dithiothreitol	1 M	7.71 g
H ₂ O	n/a	To 50 mL

Add ultrapure water to a final volume of 50 mL. Mix well and store in 1 mL aliquots at -20 °C indefinitely.

Laboratory supplies

1. 60 mL gravity column (Beyotime, catalog number: FCL60)
2. 3 K dialysis tubes (Solarbio, catalog number: YA1078)
3. 3 KDa concentration tubes (Millipore, catalog number: UFC9030)

Equipment

1. Autoclave (Shanghai Boxun Medical Biological Instrument Corp., model: YXQ-50SII)
2. Temperature-controlled incubator (Mingtu Machinery Equipment Co., model: SPX-2508)
3. Shaker incubator (Shanghai Zhichu Instrument Co., Ltd., model: ZQZY-AF8)
4. Spectrophotometer (Yipu Instrument Manufacturing Co., Ltd., model: U-T6)
5. Normal temperature, low-speed centrifuge (Xicheng Xinrui Instrument Plant, Jintan District, model: TDL-5A)
6. Freezer (-80 °C) (Thermo Fisher Scientific, model: 902ULTS)
7. Ultrasonic cell disruptor [Gallop Instrument (Shanghai) Co., Ltd, model: UCD-650B]
8. High-speed refrigerated centrifuge (Hunan Kecheng Instrument & Equipment Co., Ltd., model: W3-16R)
9. Low-speed refrigerated centrifuge (Hunan Kecheng Instrument & Equipment Co., Ltd., model: L4-5KR)
10. Electrophoresis apparatus (Shanghai Tanon Life Science Co., Ltd., model: VE-80)
11. Gel imaging system (MIULAB, model: GIS-630)
12. Singlewavelength protein purification system (GE Healthcare, model: ÄKTA prime Plus)
13. Micro high-speed refrigerated centrifuge [Gallop Instrument (Shanghai) Co., Ltd, model: MH-16KR]

Software and datasets

1. GraphPad Prism (for analysis and drawing of size exclusion chromatography)
2. ImageJ (for analysis of electrophoretic mobility shift assay)

Procedure

A. Cell transformation

1. *E. coli* BL21(DE3) cells and the expression plasmid (hSOX2-HMG-pET22b, see Figure 1) were extracted on ice.

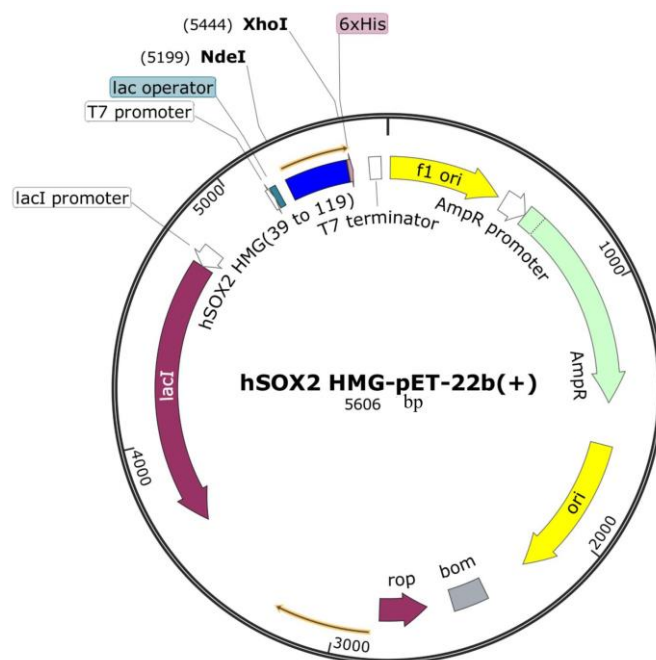


Figure 1. Amino acid sequence for hSOX2-HMG domain (Uniprot number: P48431, aa 39 to 119). MDRVKRPMNAFMVWSRGQRRKMAQENPKMHNSEISKRLGAEWKLLSETEKRPFIDEAKRLRALHMKEHPDYK YRPRRKTKTL

2. Add 1 μL of plasmid (50 ng/ μL) to the competent cells, mix gently, and incubate on ice for 5 min.
3. Heat shock at 42 $^{\circ}\text{C}$ for 1 min in a water bath, then immediately place back on ice for 2 min.
4. Recover by incubating at 37 $^{\circ}\text{C}$ in a shaker for 40 min.
5. Flame a glass spreader three times, then cool for 1 min.
6. Spread the transformed bacteria onto an LB agar plate with ampicillin (100 $\mu\text{g}/\text{mL}$, 1:1,000 dilution) using the glass spreader.
7. Close and label the plate, then incubate at 37 $^{\circ}\text{C}$ for 12–16 h.

B. Selection of positive clones

1. The next day, prepare 700 mL of LB medium, clean the workbench, and ignite an alcohol lamp.
2. Add ampicillin (final concentration 50 $\mu\text{g}/\text{mL}$, 1:2,000 dilution), mix well, and transfer 4 mL of liquid culture to the plate for scraping; then, transfer the bacterial suspension into the 700 mL LB bottle.

C. Protein expression

1. Incubate in a 37 $^{\circ}\text{C}$ shaker until OD_{600} reaches ~ 1.0 , then proceed to the second expansion.
2. Prepare two 700 mL LB cultures, add 350 μL of 100 mg/mL ampicillin (final concentration 50 $\mu\text{g}/\text{mL}$, 1:2,000 dilution), combine them with the previous culture, mix well, and divide into three bottles.
3. When OD_{600} reaches 1.2–1.5, induce expression with 0.2 mM IPTG for 4–5 h.
4. When OD_{600} reaches about 3.6, harvest cells by centrifugation at $3,072\times g$ for 15 min at room temperature and discard the supernatant.

D. Protein purification

1. Cell disruption by sonication: lyse cells (~ 40 mL) via sonication (3 s on, 5 s off) for a total of 40 min in the ice-water mixture. Take a 40 μL whole protein sample and mix with 10 μL of 5 \times SDS-PAGE loading buffer.
2. Centrifuge at $16,227\times g$ for 1 h at 4 $^{\circ}\text{C}$ and collect the supernatant. Take a 40 μL supernatant sample (contains hSOX2-HMG proteins) and mix with 10 μL of 5 \times SDS-PAGE loading buffer. See Figure 2.

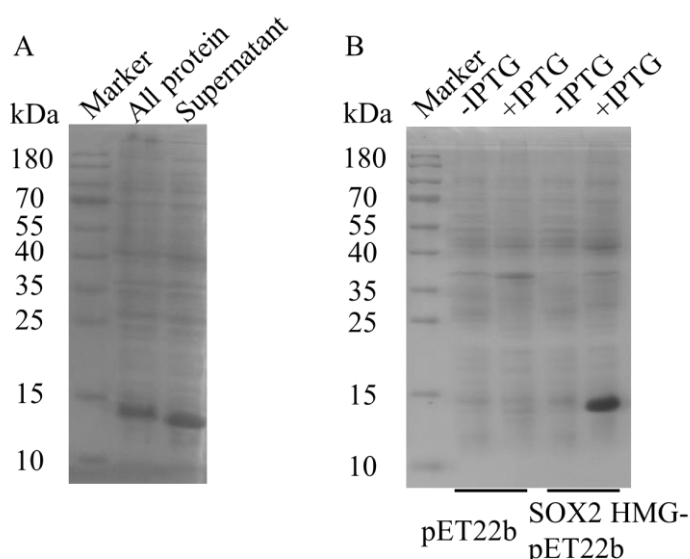


Figure 2. *E. coli* BL21(DE3) cells were lysed by sonication. Load 5 μL each of all protein and supernatant samples. Run the gel at 100 V for ~ 20 min, then increase to 160 V and run for ~ 40 min. Following electrophoresis, transfer the SDS-PAGE gel into preheated Coomassie brilliant blue G-250 and incubate for 40 min to 1 h on a rocking shaker (25 rpm). Rinse the gel thoroughly with water. To destain, add approximately 100 mL of water, heat in a microwave, and repeat the heating several times until the background is sufficiently lightened. The figure illustrates the distribution of hSOX2-HMG protein between all lysate and supernatant fractions following sonication. The majority of hSOX2-HMG protein remains soluble in

the supernatant, indicating its suitability for subsequent Ni-NTA affinity purification. The samples were analyzed by 15% SDS-PAGE gel.

3. Prepare 2 mL of Ni-NTA agarose beads. Mix the 40 mL supernatant (contains hSOX2-HMG proteins) with Ni-NTA affinity beads and incubate at 4 °C for 2–4 h with continuous mixing at 200 rpm using a magnetic stirrer.
4. Load the protein-bound Ni-NTA beads onto a pre-prepared gravity column (wash the 60 mL gravity column three times and then equilibrate with buffer) and collect the flowthrough. Close the valve to stop wash buffer flow when approximately 100 µL of wash solution remains in the column (see Figure 3).

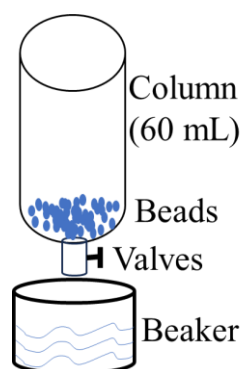


Figure 3. Schematic diagram of the purification device. The protein solution was mixed and incubated with Ni-NTA beads and placed in a 60 mL column. A beaker was used at the outlet to collect the liquid that flowed through it, and the Ni-NTA beads were left in the column without going into the beaker.

5. Wash impurities with 10 mM imidazole (100 mL, 10 washes of 10 mL each) and washing buffer 1 twice. Close the valve to stop washing buffer flow when approximately 100 µL of washing solution remains in the column.
6. Wash with 25 mM imidazole (100 mL, 10 washes of 10 mL each) and washing buffer 2 twice. Close the valve to stop washing buffer flow when approximately 100 µL of washing solution remains in the column.
7. Wash with 50 mM imidazole (50 mL, 5 washes of 10 mL each) and washing buffer 3. Close the valve to stop washing buffer flow when approximately 100 µL of washing solution remains in the column.
8. Elute with 250 mM imidazole (50 mL, 5 washes of 10 mL each) and elution buffer 1. Close the valve to stop washing buffer flow when approximately 100 µL of washing solution remains in the column. See Figure 4.

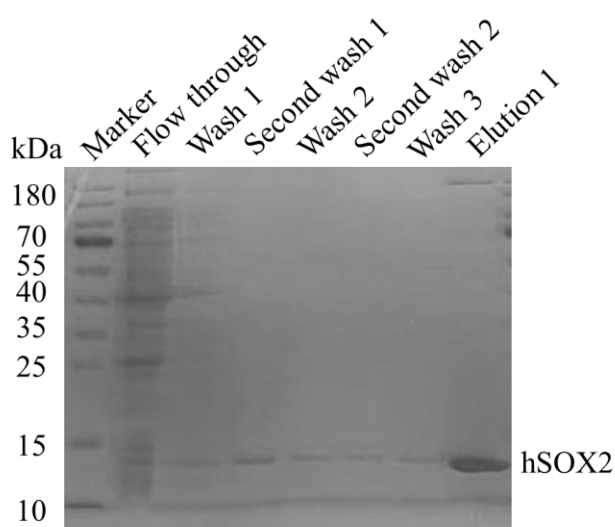


Figure 4. Purification of hSOX2-HMG protein using a detergent gradient approach. hSOX2-HMG protein purification was performed at different concentrations of imidazole buffer, as shown in the figure. Collect the hSOX2-HMG protein eluted from elution buffer 1.

E. Wash Ni-NTA beads

1. Ni-NTA beads: Wash with 50 mL of elution buffer 2 (10 mL each of 5 washes), followed by sufficient water washes.
2. Ni-NTA beads: Wash with 10 mL of 6 M guanidine, followed by sufficient water washes.
3. Store Ni-NTA beads in 20% ethanol for future use.
4. Prepare an ultrafiltration tube (cutoff 3 or 10 kDa) and 50 mL centrifuge tubes. Collect two flowthrough fractions and concentrate the protein solution to 10 mL. Divide the solution into 10 tubes, with 1 mL in each tube.
5. Pre-cool the centrifuge and centrifuge at $15,490\times g$ for 1 h at 4 °C.
6. Install the HiLoad™ 16/600 Superdex™ 200 pg column (column volume: 120 mL) onto the ÄKTA prime Plus. Equilibrate the column with 160 mL of buffer and then set up the elution program. After the column is equilibrated, rinse the 5 mL sample loop with 20 mL of H₂O. Then, equilibrate the sample loop with 20 mL of gel filtration buffer. Inject the protein solution into the sample loop and initiate the ÄKTA Prime Plus program. After the program is completed, collect protein fractions based on the chromatogram peaks and analyze by SDS-PAGE (see Figure 5).

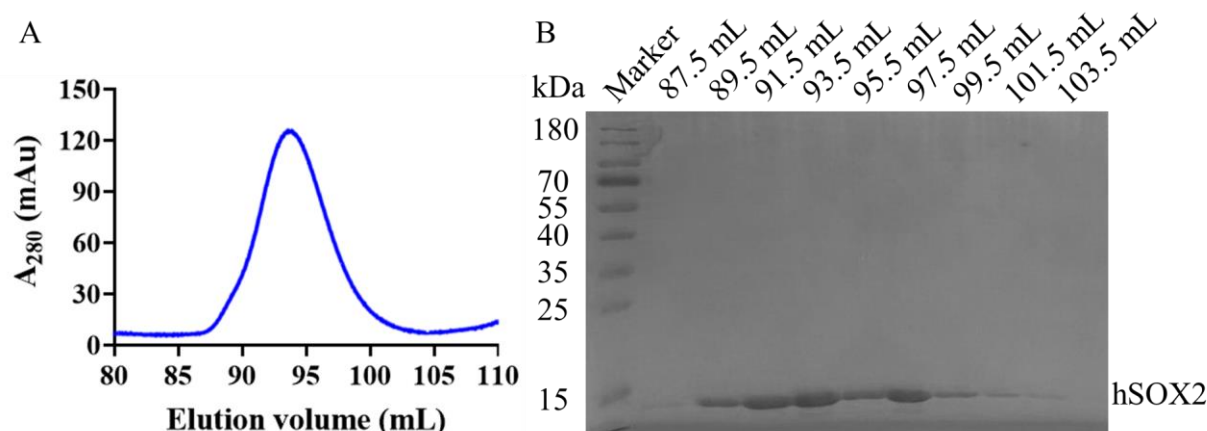


Figure 5. Gel filtration profiles and SDS-PAGE of hSOX2-HMG. (A) Size exclusion chromatography purification profile on a HiLoad™ 16/600 Superdex™ 200 pg. (B) 15% SDS-PAGE gel, showing the hSOX2 in representative elution fractions from the ÄKTA prime Plus. Lane 2 is the 87.5 mL of gel filtration, Lane 3 is the 89.5 mL of gel filtration, Lane 4 is the 91.5 mL of gel filtration, Lane 5 is the 93.5 mL of gel filtration, Lane 6 is the 95.5 mL of gel filtration, Lane 7 is the 97.5 mL of gel filtration, Lane 8 is the 99.5 mL of gel filtration, Lane 9 is the 101.5 mL of gel filtration, and Lane 10 is the 103.5 mL of gel filtration. hSOX2-HMG was purified by molecular sieves to obtain a single band, with 66.7% yield after all purification steps (Table 1). The protein concentration was measured using the BCA Protein Assay Kit. The protocol is available at https://www.biosharp.cn/index/product/details/language/cn/product_id/5268.html. Concentrations of 0.1, 0.2, 0.3, 0.4, 0.5 mg/mL BSA protein were prepared, and a standard curve was established according to the BCA Protein Assay kit protocol.

Table 1. Protein purification parameters

Protein name	Total protein volume (mL)	Whole protein concentration (mg/mL)	Total protein yield (mg)	Protein recovery yield (mg)	Recovery rate (%)
hSOX2-HMG	250	0.024	6	4	66.7

7. Electrophoretic mobility shift assay (EMSA): After gel filtration, concentrate the protein (50 μM). Centrifuge the sample at $15,490\times g$ for 5 min at 4 °C. Dilute the protein using a buffer containing 75 mM NaCl and 10 mM Tris HCl pH 7.2 to the following concentrations: 2.5 μM, 5 μM, 7.5 μM, 10 μM, 12.5 μM, and 15 μM. Prepare the target recognition sequences (double-stranded DNA) by annealing: dilute the synthesized primers P1 and P2 with ddH₂O to a concentration of 50 μM. Then, take 25 μL of P1 and P2, add them to PCR tubes, centrifuge to mix, and anneal using a PCR according to the program in Table 2. Then, dilute the target recognition sequences to a concentration of 7 μM. Set up the binding reaction according to the system described in Table 3. Add each sample sequentially and blow evenly 10 times after each addition. The reaction buffer consisted of 75 mM NaCl, 10 mM Tris-HCl pH 7.2, and 5% glycerol. Incubate the reaction mixture on ice for 2 min, followed by centrifugation at $90\times g$ for 30 s at 4 °C. Analyze the samples using polyacrylamide gel electrophoresis (PAGE) (see Figure 6A). Assuming no aggregation, the EMSA results are analyzed [5], which is briefly described as a quantification

of each of the three sets of parallel EMSA results by ImageJ. The relative fraction of DNA was obtained by pre-processing the PAGE gel plots using ImageJ, then framing the bands of free DNA, and calculating the DNA area of each lane to obtain the relative fraction of DNA, with [1 - free DNA fraction] representing the bound fraction. Binding data were fitted with the Hill equation and analyzed in GraphPad to determine apparent K_d values.

Table 2. Annealing program

	Temperature	Time
1	98 °C	5 min
2	98 °C	5 min
Increase	-5 °C/cycle	14 cycle
3	GOTO 2	

Table 3. Reaction solution of EMSA

Ratio	0	0.25	0.50	0.75	1.00	1.25	1.50
Reaction buffer	15 µL	13 µL	13 µL	13 µL	13 µL	13 µL	13 µL
BSA	2 µL	2 µL	2 µL	2 µL	2 µL	2 µL	2 µL
DNA	3 µL	3 µL	3 µL	3 µL	3 µL	3 µL	3 µL
Protein	0 µL	2 µL	2 µL	2 µL	2 µL	2 µL	2 µL

Note: Table 3 shows the volume of different samples added. 10 µM BSA was prepared, and DNA was prepared to 7 µM. Different concentrations of 2.5, 5, 7.5, 10, 12.5, and 15 µM hSOX2-HMG protein were prepared. The ratios (0, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50) correspond to the molar ratio of protein to DNA in the 20 µL binding reaction.

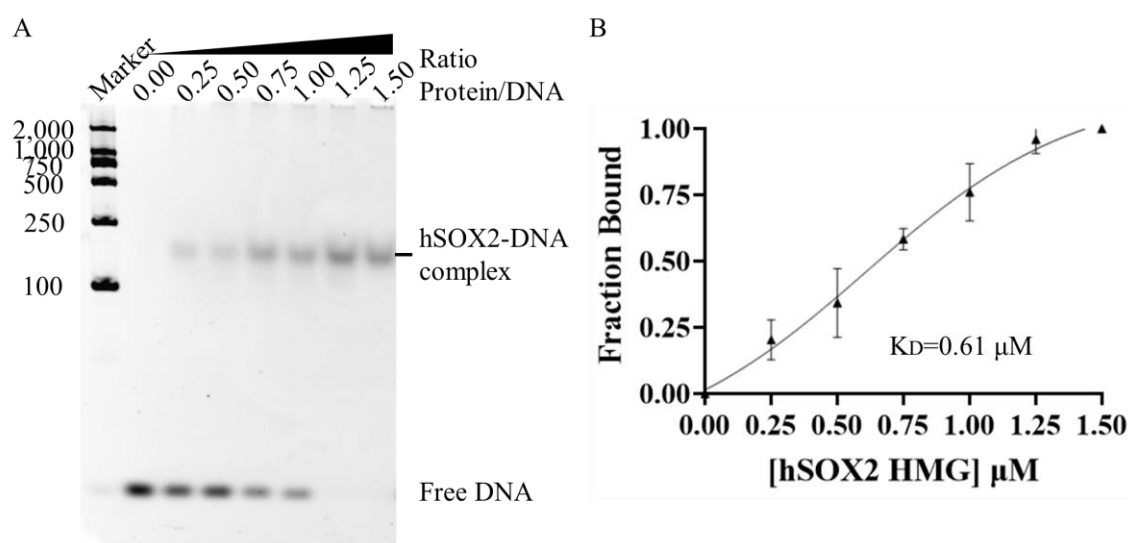


Figure 6. hSOX2-HMG protein activity analysis. (A) Electrophoretic mobility shift assay of hSOX2-HMG-DNA interaction. A representative image is shown for three groups of parallel experiments. 0.00 ratio: 0.00 µM SOX2-HMG to 1 µM DNA. 0.25 ratio: 0.25 µM SOX2-HMG to 1 µM DNA. 0.50 ratio: 0.50 µM SOX2-HMG to 1 µM DNA. 0.75 ratio: 0.75 µM SOX2-HMG to 1 µM DNA. 1.00 ratio: 1 µM SOX2-HMG to 1 µM DNA. 1.25 ratio: 1.25 µM SOX2-HMG to 1 µM DNA. 1.50 ratio: 1.50 µM SOX2-HMG to 1 µM DNA. (B) Quantification of the bound fraction using the results shown in (A). When the ratio is 0.00, 0.25, 0.50, 0.75, 1.00, 1.25, and 1.50, the Fraction Bound-values are 0.00 ± 0.00 , 0.20 ± 0.07 , 0.34 ± 0.13 , 0.58 ± 0.04 , 0.76 ± 0.11 , 0.96 ± 0.05 , and 1.00 ± 0.00 , respectively. Error bars represent standard deviation values from three groups of parallel experiments.

Validation of protocol

1. We performed at least three expression and purification cycles of hSOX2-HMG, and different people confirmed the activity of the prepared hSOX2-HMG protein through EMSA. Our data indicates that the protocol is robust and reproducible.

The results showed that protein purity was greater than or equal to 90% (Figure 5B and Figure 7A) after standardization and 1 L expression. Through the purification, we were able to obtain 2–2.86 mg of hSOX2-HMG protein (Table 1 and Figure 7C). The protein concentration of each sample was determined by relating its absorbance value to the standard curve (BCA standard curve: $Y = 2.211 \times X - 0.006$, Figure 7B).

Additionally, EMSA results showed no free DNA at ratios of 1.25 or 1.50, indicating that all DNA formed a complex with hSOX2-HMG recombinant protein. The EMSA analysis revealed a hSOX2–DNA complex migrating between 100 and 250 bp in the native PAGE gel (Figure 6A), with a dissociation equilibrium constant (K_d) of $\sim 0.61 \mu\text{M}$ (Figure 6B). The results confirm that the hSOX2-HMG recombinant protein obtained through the described purification method retains its DNA-binding activity.

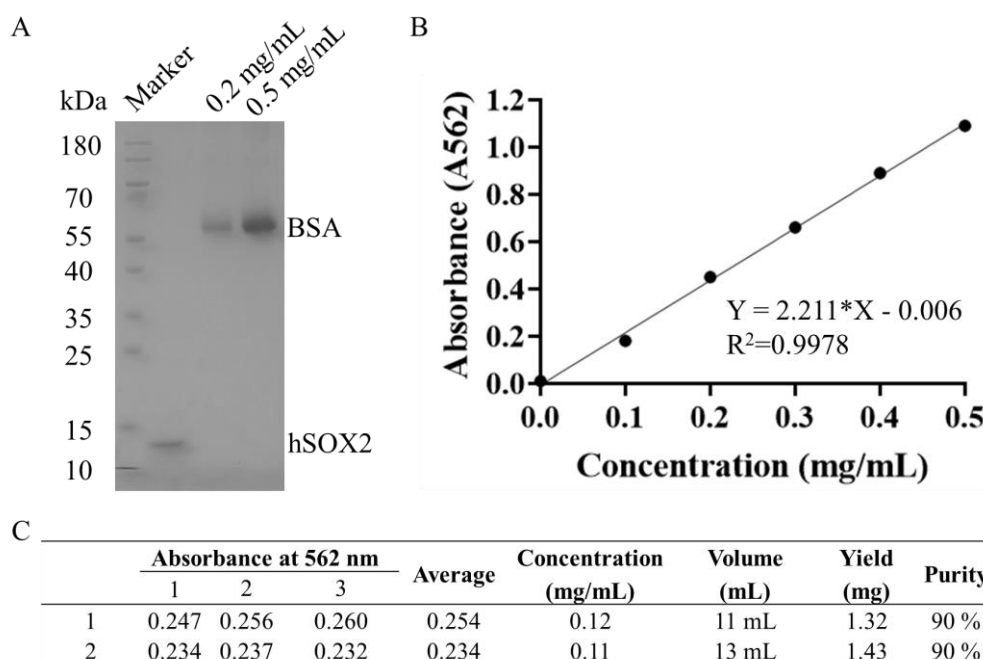


Figure 7. Concentration quantification and purity analysis of hSOX2-HMG. (A) Representative SDS-PAGE of purified recombinant protein after the purification of gel filtration. A representative image is shown for two parallel experiments. Standard control for protein concentration: 0.2 mg/mL and 0.5 mg/mL BSA. (B) BCA standard curve. (C) Concentration and yield quantification for hSOX2-HMG of a 500 mL cell culture in which recombinant protein of hSOX2-HMG was induced and subsequently purified by Ni-NTA and gel filtration, as described above in this protocol.

Acknowledgments

This protocol was done under the conception and guidance of supervisor Professor Jingjun Hong. We thank the core facility of the Institute of Health Sciences and Technology (IHST) in Anhui University.

Funding: This work was supported by National Natural Science Foundation of China (NSFC, 31970669), Double First Class Construction Funds (S020318006/067), and College Student Innovation and Entrepreneurship Training Program Project (S202410357310).

The following figures were created using BioRender: Graphical overview, created in BioRender. Hong, J. (2025) <https://BioRender.com/wgh1dti>.

Competing interests

The authors declare no competing interests.

Received: April 10, 2025; Accepted: June 22, 2025; Available online: August 7, 2025; Published: August 20, 2025

References

1. Kamachi, Y. and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. *Development*. 140(20): 4129–4144. <https://doi.org/10.1242/dev.091793>
2. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature*. 346(6281): 245–250. <https://doi.org/10.1038/346245a0>
3. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*. 131(5): 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>
4. Morales, L., Hernández, P. and Chaparro-Olaya, J. (2018). Systematic Comparison of Strategies to Achieve Soluble Expression of Plasmodium falciparum Recombinant Proteins in *E. coli*. *Mol Biotechnol*. 60(12): 887–900. <https://doi.org/10.1007/s12033-018-0125-0>
5. Chittori S, Hong J, Saunders H, Feng H, Ghirlando R, Kelly AE, Bai Y, Subramaniam S. (2018). Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science*. 359(6373):339-343. <https://doi.org/10.1126/science.aar2781>