

**Protein Expression and Purification of Integrin I Domains and IgSF Ligands
for Crystallography**

Hongmin Zhang* and Jia-huai Wang#

*Department of Physiology, the University of Hong Kong, Hong Kong SAR, China

Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115

Abstract

Cell adhesion depends on combinational expression and interactions of a large number of adhesion molecules from opposing cells. Integrins and immunoglobulin superfamily (IgSF) members are two foremost classes of cell adhesion molecules in immune system.

Structural study is critical for better understanding of the interactions between integrins and their IgSF ligands. Here we describe protocols for protein expression of integrin α L I domain and its IgSF ligand ICAM-5 D1D2 fragment for crystallography.

Key words: mammalian cell expression; integrins; IgSF ligands; ICAMs; crystallization; molecular replacement.

1. Introduction

Integrins are major cell adhesion molecules that mediate cell-cell and cell-extracellular matrix interactions, thereby playing a key role in development, immune responses, leukocyte trafficking, homeostasis and cancer metastasis. Integrins transduce signals across the plasma membrane bi-directionally in an allosteric fashion. Ligand-binding to integrins transmits signals to the cytoplasm (“outside-in” signaling). Conversely, integrins can be activated in response to intracellular signaling cascades elicited by other receptors (“inside-out” signaling) (1, 2). The structural basis of integrin allostery has been extensively reviewed (3). Structural studies of the binding domain of leukocyte integrins (α I domain) and their ligands have been extensively performed to understand the allosteric regulation of integrins. For example, the complex structures between integrin α L I domain and its ligands ICAM-1, ICAM-3 and ICAM-5 (4-6) have revealed a basic binding model between integrins and their ligands. An acidic residue from the ligand coordinates to the metal ion dependent adhesion site (MIDAS) of I domain (7), thereby triggering the conformational changes of MIDAS, which is allosterically linked to an axial movement of the α 7-helix at the other end of the I domain. This eventually leads to a large-scale reorientation of the ecto-domains up to 200Å, and the separation of the integrin α and β subunits by as much as 70Å (8, 9). In this chapter, we will use the α I domain of α L β 2 integrin and its natural ligands, ICAM-5 as an examples to describe the expression and structure determination of integrin domains in complex with ligand.

2. Materials

2.1 Expression and purification of α L I domain in *E.coli*

1. Expression vector: pET22b with an inserted fragment encoding the α L I domain (residue N129 to Y307 with a stop codon following Y307) (**10**, **11**).
2. BL21 (DE3) competent cells.
3. LB-amp medium: Dissolve 10g tryptone, 5g yeast extract and 10g sodium chloride in 1L of water (*see Note 1*). Autoclave at 121°C for 20min, add 1ml of 100 mg/ml ampicillin when it is cooled to room temperature.
4. Rich medium: 20 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, 20 ml/L glycerol, 50 mM K_2HPO_4 , 10 mM $MgCl_2$, 10 g/L glucose, and 100 μ g/L ampicillin.
5. 0.1 M Isopropyl β -D-1-thiogalactopyranoside (IPTG).
6. Lysis buffer: 50 mM Tris-HCl pH8.0, 1 mM $MgCl_2$, 0.4 μ g/ml DNase I, 0.4 μ g/ml RNase A, 1 μ g/ml lysozyme.
7. Sonicator with a large sonication tip.
8. Wash buffer 1: 20 mM Tris-HCl pH8.0, 23 % (w/v) sucrose, 0.5 % (v/v) Triton X-100, 1 mM EDTA.
9. Wash buffer 2: 20 mM Tris-HCl pH8.0, 1 mM EDTA.
10. Solubilization buffer: 6 M Guanidine HCl, 50 mM Tris-HCl pH8.0, 1 mM DTT.
11. Refolding buffer: 50 mM Tris-HCl pH8.0, 1 mM $MgCl_2$, 5 % (v/v) glycerol, 50 mg/L $CuSO_4$ (*see Note 2*), 1 mM phenanthroline (stock of 1M in DMSO), 0.1mM PMSF (stock of 100 mM in isopropanol).

12. Ion exchange buffer A: 20 mM Tris-HCl pH8.0.
13. Ion exchange buffer B: 20 mM Tris-HCl pH8.0, 1 M sodium chloride.
14. Size exclusion buffer: 20 mM HEPES pH7.5, 0.2 M sodium chloride.
15. FPLC equipment.
16. Q-sepharose ion-exchange column.
17. Superdex75prep size-exclusion column.

2.2 Expression and purification of ligands in CHO-*lec* 3.2.8.1 cells

1. CHO *lec* 3.2.8.1 cells stably transfected to express ICAM-5 D1D2 (6).
2. GMEM-MSX medium: To make 500 ml or 2L of GMEM-MSX medium, add the following stock solutions as listed below using aseptic technique in a cell culture hood.

Sterile water (ml)	350	1400
a. 10XGMEM (ml)	50	200
b. Sodium bicarbonate (ml)	18.1	72.4
c. NEAA (ml)	5	20
d. G+A (ml)	5	20
e. Sodium pyruvate (ml)	5	20
f. Nucleosides (ml)	10	40
g. Pen-Step (ml)	5	20
h. Dialyzed FCS (ml)	50	200

i. L-MSX (ml)	0.125	0.5
Total (ml)	500	2000

Stock solutions:

- a. 10X Glasgow MEM without glutamine and without tryptose-phosphate broth (custom order to GIBCO or Sigma).
- b. 7.5% sodium bicarbonate.
- c. 100X Non-essential amino acids (NEAA).
- d. 100X glutamic acid + asparagine (G+A): Dissolve 1,500 mg L-glutamic acid and 1,500 mg L-asparagine to 250 ml of distilled water and sterilize by filtration.
- e. 100 mM sodium pyruvate.
- f. 50X Nucleosides: Dissolve 175 mg adenosine, 175 mg guanosine, 175 mg cytidine, 175 mg uridine, 40 mg thymidine to 500 ml of autoclaved water, and sterilize by filtration.
- g. 100X Penicillin-Streptomycin at 5,000units/ml.
- h. Dialyzed FCS: Heat inactivated at 56 °C for 30~35 min (*see Note 3*).
- i. 100 mM L-MSX: Prepare 18 mg/ml solution in PBS. Sterilize by filtration and store at -20°C in 1-ml aliquots. Final concentration in medium is 25 µM.

4. FPLC equipment.
5. Cation exchange column SP.
6. Anion exchange column Mono Q.
7. Cation exchange column Mono S.
8. Size exclusion column Superdex75 prep.

3. Methods

The α L I domain is expressed in *E.coli* as an inclusion body, refolded, and purified to homogeneity. By contrast, the integrin ligand domain of ICAM-5 (ICAM-5D1D2) is expressed in CHO cells, as the protein is highly glycosylated and involves several disulfide bonds. A mutant cell line CHO *lec* 3.2.8.1 has four independent mutations in the N- and O-glycosylation pathways (*12*). N-linked carbohydrates produced by CHO *lec* 3.2.8.1 cells are all of the high mannose type, but different in the number of mannoses, ranging from Man₅ to Man₉. O-glycosylation is homogenous, with only a single GalNAc residue attached per site. When cultured in the presence of the alpha-glucosidase I inhibitor N-butyl-deoxynojirimycin (NB-DNJ), glycoproteins produced in CHO *lec* 3.2.8.1 cells are almost completely susceptible to Endo H digestion (*13, 14*). Endo H cleaves chitobiose, leaving a single N-linked N-acetylglucosamine per site, which is ideal for maintenance of protein solubility and special carbohydrate-protein interactions, such as between the first N-acetyl glucosamine residue and tryptophan. The property of CHO *lec* 3.2.8.1 makes it quite suitable for expression of proteins for structural studies. Therefore, CHO *lec* 3.2.8.1

will be our expression host for integrin IgSF ligands including ICAM-5 D1D2.

3.1 Expression and purification of α L I domain in *E.coli*

1. Inoculate a single colony of BL21(DE3) transformed with the expression vector into 30 ml of LB-ampicillin medium and shake at 37 °C for 6~8 hours.
2. The next day, transfer 30 ml of bacteria culture to 600 ml of rich medium, shake at 37 °C for 3~4 hours. When OD600 reaches to 1~1.2, add 1 mM IPTG to induce expression.
3. 3~4 hours later harvest bacteria by centrifugation. Re-suspend the pellet with 50 mM Tris-HCl pH8.0 and centrifuge to harvest the pellet. Freeze pellet at -20°C for later use.
4. Suspend the pellet in 30 ml of lysis buffer, incubate it at 37°C for 10~15min, and sonicate it with a large sonicator tip.
5. Harvest inclusion bodies by centrifuge (25,000 X g for 30min).
6. Wash inclusion bodies with wash buffer 1, centrifuge at 25,000 X g for 15min and discard the supernatant. Repeat this step for 5 times.
7. Wash inclusion bodies with wash buffer 2.
8. Resuspend inclusion bodies in solubilization buffer, stir at room temperature for 1~2 hours.
9. Centrifuge at 25,000 X g for 30 min, filter and keep the supernatant. Measure the protein concentration at OD280 and adjust protein concentration to 0.5~1 mg/ml

with solubilization buffer.

10. Refolding at 4°C overnight or longer by quick dilution into 19-fold of refolding buffer.
11. Centrifuge and filter with a 0.22 µm membrane to remove precipitant.
12. Concentrate the supernatant to small volume and filter it.
13. Purify the sample on a Q-sepharose ion-exchange column using a FPLC equipment. The supernatant is loaded to the column, washed with 2 column volumes (CV) of buffer A and eluted with a NaCl gradient from 0 to 30% of buffer B in 20 CV.
14. Collect and pool the peak fractions.
15. Concentrate the pooled fractions and further purify it on Superdex75 prep size-exclusion column using a FPLC equipment.
16. Collect and pool the peak fractions.
17. Desalt and change buffer to 20 mM HEPES pH7.5, 50 mM sodium chloride and 5 mM magnesium chloride.
18. Concentrate the sample to >20 mg/ml. Flash freeze into liquid nitrogen and store at -80 °C for later use of crystallization (*see Notes 5*).

3.2 Expression and purification of ligands in CHO-lec 3.2.8.1 cells

Our ICAM-5D1D2 construct does not contain any purification tags and we have no suitable antibodies against ICAM-5D1D2 to be used in affinity purification. However, the

unique situation here is that the isoelectric point (pI) of ICAM-5 D1D2 is about 11, far above the pI value of most proteins (4~6). This property was exploited to purify this protein fragment using a series of ion-exchange columns.

1. Collect supernatant from CHO *lec* 3.2.8.1 cells stably transfected to express ICAM-5 D1D2. The cells were cultured in dishes, flasks or in roller bottles with GMEM-MSX medium. Roller bottles generally give higher expression due to larger surface area. We typically harvest the supernatant every week and add fresh medium into roller bottles (100~200 ml for a 2-L roller bottle) for long term culture (*see Note 4*).
2. Concentrate the supernatant up to 10~20 folds and dialyze it against 50 mM Tris-HCl, pH8.8.
3. Centrifuge to remove precipitants.
4. Load the concentrated supernatant to a SP cation exchange column pre-equilibrated with 50 mM Tris-HCl pH8.8. Wash the column with 10 CV of 50 mM Tris-HCl pH8.8. Elute ICAM-5D1D2 with 50 mM Tris-HCl pH8.8, 1M NaCl.
5. Collect and pool peak fractions containing ICAM-5D1D2. Dialyze it against 50 mM Tris-HCl pH8.8.
6. Load the dialyzed sample to Mono Q anion exchange column pre-equilibrated with 50 mM Tris-HCl pH8.8. ICAM-5D1D2 does not bind to the column at this pH. Collect the flow through fraction. Although ICAM-5D1D2 does not bind to Mono Q at this pH, contaminant proteins do and this step improves the purity.

7. Load the flow through fraction from Mono Q to Mono S pre-equilibrated with 50 mM Tris-HCl pH8.8. Wash the column with 2CV of 50 mM Tris-HCl pH8.8 and elute with a NaCl gradient from 0 to 0.4 M in 20 CV. Collect and pool the peak fractions.
8. Concentrate the ICAM-5D1D2 fractions from Mono S and load to Superdex 75 prep column pre-equilibrated with 20 mM HEPES pH7.5, 0.2 M NaCl. Collect and pool the peak fractions.
9. Concentrate the pooled fractions and change buffer by dialysis to 20 mM HEPES, pH7.5, 50 mM NaCl. Aliquote, flash freeze into liquid nitrogen and store at -80°C for later use (*see* **Notes 5** and **6**).

4. Notes

1. Unless stated otherwise, all solutions should be prepared in water that has a resistivity of 18.2 M Ω -cm and total organic content of less than five parts per billion. This standard is referred to as “water” in this text.
2. For wild type I domain, CuSO₄ and phenanthroline should not be included. CuSO₄ is used to facilitate oxidization of the engineered disulfide-bond in the mutant locked high-affinity I domains and phenanthroline is used to inhibit metalloprotease activity. GSH and GSSH may be used instead of CuSO₄ to provide oxidization-reduction potential. CuSO₄ and phenanthroline are less expensive.
3. Dialyzed FCS should be used especially when culturing CHO cells in the presence

of L-MSX.

4. CO₂ is not needed for CHO *lec* cells cultured in roller bottles. The bottles can be placed on a rotating rack at a speed of 2~3 rpm in a 37°C warm room. Harvest the supernatant and add in fresh medium with an aseptic technique in a tissue culture hood. Parafilm can be used to wrap around the bottle cap to further reduce contaminant.
5. In the crystallization of ICAM-5 D1D2 with the integrin α_L I domain, ICAM-5 D1D2 was mixed with the I domain in equal molar ratio at a total concentration of 11 mg/ml. The mixture was used in crystallization screening and we obtained crystals from precipitant in a reservoir solution containing 0.1 M HEPES pH 7.5, 10% PEG8000, 8% ethylene glycol at room temperature. Crystals were optimized with a pH gradient from 7.0 to 8.0 and PEG8000 gradient from 5% to 15%. Later, ethylene glycol was replaced by glycerol with a gradient from 5% to 12%. Better crystals were obtained with 2 μ l of protein mixed with 1 μ l of reservoir solution (0.1 M HEPES pH 7.5, 7.5% PEG8000, 10% glycerol) at 4° C in 3 days.

The crystals were harvested and soaked in 0.1M HEPES pH7.5, 15% PEG8000, 20% glycerol and 5 mM magnesium chloride. It is important to increase the concentration of PEG8000 from 7.5% to 15%, otherwise the crystals dissolved slowly and this impairs the diffraction quality. After soaking in cryo-protectant, the crystals were flash frozen into liquid nitrogen for later data collection.

6. The diffraction data for crystals of ICAM-5D1D2 in complex with the integrin α_L I

domain were collected at ID19 at Argonne National Laboratory and processed with HKL2000 (*15*). The scaled diffraction data (sca file) were fed into program dtrek2mtz in CCP4 (*16*) and converted into mtz file with 5% of the data added to FreeR column. The result of Mathew's coefficient calculation (*17*) showed that there was likely only one copy of the complex in the asymmetric unit with about 67% solution content. Self-rotation function did not show any pseudo-symmetry, and there was no pseudo-translation either. Thus, we figured that there should be only one copy of the complex in the asymmetric unit to look for.

Two homologous structures, ICAM-3 D1 in complex with high affinity I domain (HA) (PDB code 1T0P) and ICAM-1 D1D2 with intermediate affinity I domain (IA) (PDB code 1MQ8), were used as search models to solve the structure of ICAM-5D1D2 with I domain by molecular replacement. The structure of ICAM-1 D1D2 alone was also used (PDB code 1IC1). Phaser (*18*) from CCP4 package was used to accomplish molecular replacement with HA, ICAM-3D1 and ICAM1-D2 as search models.

The solution was refined in Refmac (*19*) in CCP4 suit by rigid body refinement followed by constrained refinement. By alternate model building with Coot (*20*) to fit density into ICAM-5 sequence and refinement with Refmac, there was not much difficulty in the tracing of ICAM-5. However, we did encounter some trouble in correctly tracing the I domain. In all previous integrin α I domain structures, $\alpha 7$ helix is located between $\beta 6$ strand and $\alpha 1$ helix with a direction from

top to down. At the beginning we tried to model the $\alpha 7$ helix of I domain in this direction and the density did not fit well with the residues of $\alpha 7$ helix. Furthermore, the $\beta 6$ - $\alpha 7$ loop did not have any reasonable density (**Fig. 1a**). When the symmetry related molecules were checked, we noticed that $\alpha 7$ helix of one molecule might swing out and insert into a symmetry related molecule in an upside-down fashion (**Fig. 1a**). **Fig. 1b** shows a typical α helix, for which all the $C\alpha$ - $C\beta$ bond point to the N terminus of the helix. Upon carefully checking the density we noticed that most of the residues in $\alpha 7$ helix showed a polarity with $\alpha 7$ helix from down to up (**Fig. 1c**). The final model did have a swing-out of the $\alpha 7$ helix (**Fig. 2**).

References

1. Hynes, R. O. (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* **110**, 673-687.
2. Springer, T. A., and Wang, J. H. (2004) The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion. *Adv. Protein Chem.* **68**, 29-63.
3. Luo, B. H., Carman, C. V., and Springer, T. A. (2007) Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.* **25**, 619-647.
4. Shimaoka, M., Xiao, T., Liu, J. H., Yang, Y., Dong, Y., Jun, C. D., McCormack, A., Zhang, R., Joachimiak, A. and Takagi, J. (2003) Structures of the alpha L I Domain and Its Complex with ICAM-1 Reveal a Shape-Shifting Pathway for Integrin Regulation. *Cell* **112**, 99-111.
5. Song, G., Yang, Y., Liu, J. H., Casasnovas, J. M., Shimaoka, M., Springer, T. A., and Wang, J. H. (2005) An atomic resolution view of ICAM recognition in a complex between the binding domains of ICAM-3 and integrin alphaLbeta2. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3366-3371.
6. Zhang H., Casasnovas J.M., Jin M., Liu J.H., Gahmberg C.G., Springer T.A. and Wang J.H. (2008) An unusual allosteric mobility of the C-terminal helix of a high-affinity alpha L integrin I domain variant bound to ICAM-5. *Mol. Cell* **31**, 432-437.
7. Lee J.O., Rieu P., Arnaout M.A. and Liddington R. (1995) Crystal structure of the

- A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* **80**, 631-638.
8. Xiao T., Takagi J., Collier B.S., Wang J.H. and Springer T.A. (2004) Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* **432**, 59-67.
 9. Xiong, J.P., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D.L., Joachimiak, A., Goodman, S.L., Arnaout, M.A. (2001) Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**, 339-345.
 10. Shimaoka M., Lu C., Palframan R.T., von Andrian U.H., McCormack A., Takagi J., Springer T.A. (2001) Reversibly locking a protein fold in an active conformation with a disulfide bond: integrin alphaL I domains with high affinity and antagonist activity in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 6009-6014.
 11. Jin M., Song G., Carman C.V., Kim Y.S., Astrof N.S., Shimaoka M., Wittrup D.K., Springer T.A. (2006) Directed evolution to probe protein allostery and integrin I domains of 200,000-fold higher affinity. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5758-5763.
 12. Stanley P. (1989) Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell. Biol.* **9**, 377-383.
 13. Davis S.J., Davies E.A., Barclay A.N., Daenke S., Bodian D.L., Jones E.Y., Stuart D.I., Butters T.D., Dwek R.A. and van der Merwe PA. (1995) Ligand binding by

- the immunoglobulin superfamily recognition molecule CD2 is glycosylation-independent. *J. Biol. Chem.* **270**, 369-375.
14. Ikemizu S., Sparks L.M., van der Merwe P.A., Harlos K., Stuart D.I., Jones E.Y. and Davis S.J. (1999) Crystal structure of the CD2-binding domain of CD58 (lymphocyte function-associated antigen 3) at 1.8-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4289-4294.
 15. Otwinowski Z. and Minor W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, Volume 276: Macromolecular Crystallography, part A, p.307-326.
 16. Collaborative Computational Project, Number 4. 1994. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* **D50**, 760-763.
 17. Matthews B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
 18. McCoy A. J., Grosse-Kunstleve R. W., Adams P. D., Winn M. D., Storoni L.C. and Read R.J. (2007). Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658-674.
 19. Murshudov G.N., Vagin A.A. and Dodson E.J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood method. *Acta. Cryst.* **D53**, 240-255.
 20. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132.

Fig. 1. Swung-out of the $\alpha 7$ helix of the I domain in complex with ICAM-5 D1D2. (A) Fo-Fc map showing the traces of $\beta 6$ - $\alpha 7$ loop and $\alpha 7$ helix. The positive map of Fo-Fc was shown as grey mesh. The I domain and ICAM-5 were shown as C α traces. The density connecting to Leu289 showed an upward tracing for $\beta 6$ - $\alpha 7$ loop and $\alpha 7$ helix. And there was a break between Leu289 and the density at the lower part of I domain. (B). A typical α helix composed of alanines. All C α -C β bonds of the helix pointed to the N terminus of the helix. (C). Polarity of the density for $\alpha 7$ helix showing a direction from down to up. The $\alpha 7$ helix was modeled into the density and was shown as C α traces. Some of the residues were labeled showing a clear polarity from down to up.

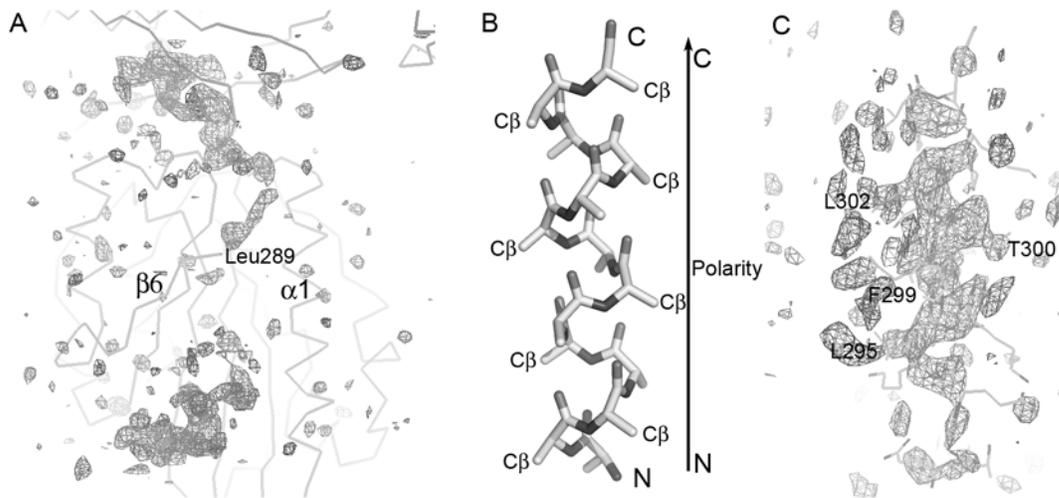


Fig. 2. Ribbon diagram of ICAM-5/I domain complex. Two symmetry-related ICAM-5/I domain complexes were shown with the C-terminal $\alpha 7$ -helix of one I domain (at the lower position) inserted into a groove of the other I domain (at the upper position) and the $\alpha 7$ -helix of upper I domain swung-out to insert into the third I domain, which was not shown in the figure.

