인공 단백질 나노구조체 최신 연구동향

한국과학기술원 이지오 교수

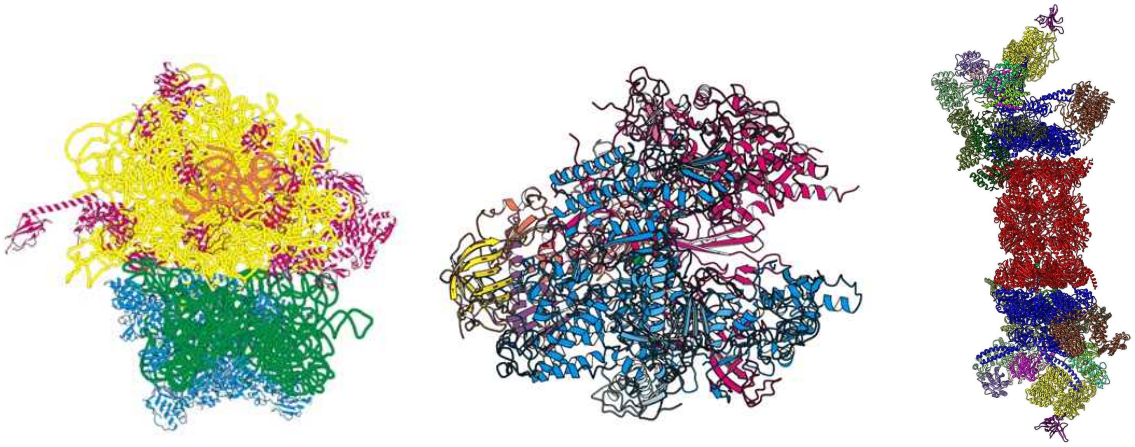
1. 개요

세포가 정상적인 기능을 수행하기 위해서는 매우 다양한 단백질 복합체가 이용되고 있다. 예를 들어 세포 기능의 핵심이 되는 전사, 복제, 번역, splicing, proteasome 기능은 모두 수십 ~ 수백 개의 단백질과 핵산이 복잡하고 다양한 복합체를 형성하여 핵심 역할을 하고 있다 (그림 1). 생물학적인 시스템에서 단백질 복합체는 다음과 같은 이유로 중요하다. (1) 시스템의 효율성. 하나의 생체 대사경로에 참여하는 서브유닛이 모여 있으면, 중간 대사물이 확산될 필요가 없기 때문에 전체 시스템의 효율성이 높아진다. (2) 조절의 용이성. 하나의 생체 대사경로에 참여하는 단백질 서브유닛이 모여 있으면, 기능을 하는 서브유닛과 조절작용을 하는 서브유닛이 서로 상호작용을 할 수 있기 때문에 보다 효율적인 조절이 가능하다. (3) 다양한 기능유닛을 모아주는 플랫폼 역할. 세포내부의 필요에 따라 여러 서브유닛들이 모이게 되면, 기존의 기능과는 다른 기능을 수행할 수 있게 된다. 여러 서브유닛을 모아주는 플랫폼 단백질은 특히 고등 생명체에서 더욱 중요한 역할을 한다. (4) 세포의 모양이나 구조를 형성하는 역할. 세포의 모양과 움직임은 세포의 기능에서 중요한 역할을 하게 되는데 세포의 모양과 모양변화는 많은 수의 단백질이 모여서 복합체를 형성하면서 이루어진다. 또한, 바이러스의 외부 껍질은 24개나 60개의 단백질 서브유닛으로 이루어져 있다. 바이러스의 단백질 껍질은 내부의 유전물질을 보호하며, 세포막에 있는 수용체와 결합하는 역할을 한다.

최근 세포에 있는 단백질 복합체와 유사한 인공 단백질 복합체를 만들려고 하는 연구가 활발히 진행되고 있다. 이러한 연구는 크게 두 개의 흐름으로 분류해 볼 수 있다. 첫째는 단백질이나 핵산 서브유닛들을 구조적인 유동성이 큰 링커로 연결하여 대형 복합체를 만드는 연구이다. 이렇게 만들어진 복합체는 일정한 구조를 가지지 못하고 다양한 구조가 섞여있는 상태가 되며, 각각의 서브유닛과는 다른 새로운 기능적 특

성을 가지는 경우가 많다. 이와 같은 stochastic protein complex 는 여러 분야에서 응용성이 연구되고 있다. 둘째는 단백질 서브유닛을 미리 예측이 가능한 정확한 구조를 가지도록 디자인하는 연구이다. 정확한 구조를 성공적으로 디자인하는 데에는 매우 정교한 설계가 필요하기 때문에 최근까지도 성공한 예가 많지 않았다. 하지만, 최근에 단백질 구조와 움직임에 대한 보다 정확한 계산이 가능해지면서, 성공적인 디자인이 발표되기 시작하고 있다. 예측가능한 단백질 나노구조를 이용하면 자연계에는 존재하지 않고, 응용성이 높은 새로운 단백질 복합체 생산이 가능해 질 것으로 기대되고 있다.

본 보고서에서는 대표적인 최근 사례들을 중심으로 인공단백질 구조체 개발 연구 동향에 대해 고찰해 보고자 한다.

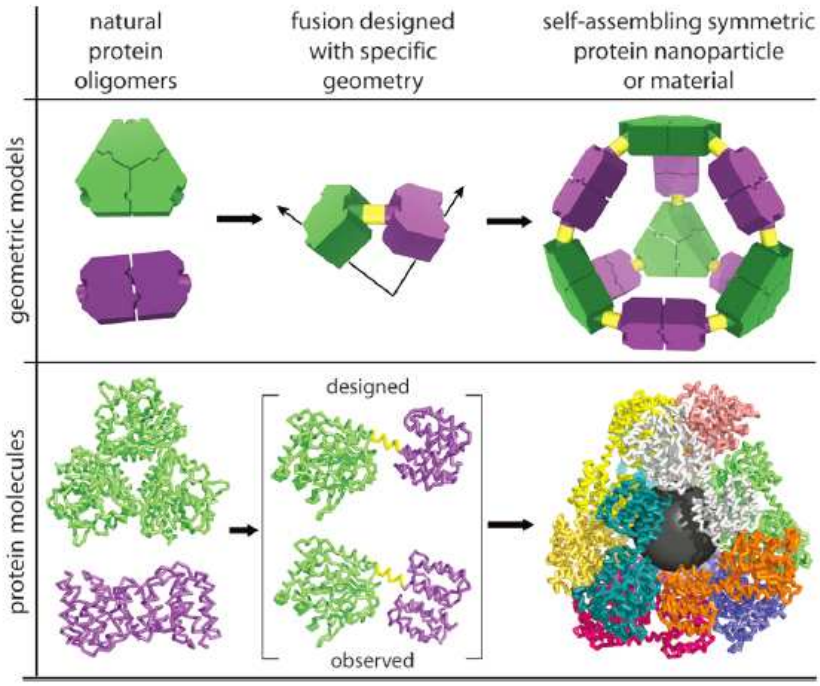


<그림 1>. 세포에 존재하는 대표적인 단백질 복합체. 왼쪽부터 리보좀, RNA polymerase, 프로테아좀 구조.

2. 알파헬릭스 링커를 이용한 케이지 단백질 디자인

알파헬릭스는 80% 이상의 단백질이 가지고 있는 이차구조이다. 알파헬릭스는 정확하게 예측이 가능한 구조를 가지고 있으며, 베타 스트랜드에 비하여 독립적인 구조를 가지고 있기 때문에, 비교적 쉽게 디자인이 가능하다. 만약 두 개의 단백질을 알파헬릭스 구조를 가지는 링커로 연결하면 두 단백질의 상대적인 구조가 정확하게 예측이 가능할 것이다. 이러한 성질을 이용하여 자연적으로 trimer를 형성하는 단백질과 dimer를 형성하는 단백질을 연결하면 다양한 단백질 구조체를 만들 수 있다 (Padilla et al., 2001). 2012년 Todd Yeates 그룹은 trimer를 형성하는 bromoperoxidase (BPO) 단백질과 dimer를 형성하는 M1 바이러스 matrix 단백질을 알파헬릭스로 연결하는데에 성공하였다. BPO는 단백질 끝부분이 알파헬릭스이고, M1 단백질은 알파헬릭스로 시작되며, 이를 연결하는 헬릭스 링커는 109.5 도 각도로 향하고 있기 때문에 tetrahedral cage 구조가 가능할 것으로 예측되었다. 연구자들은 많은 수의 링커를 테스트하였는데

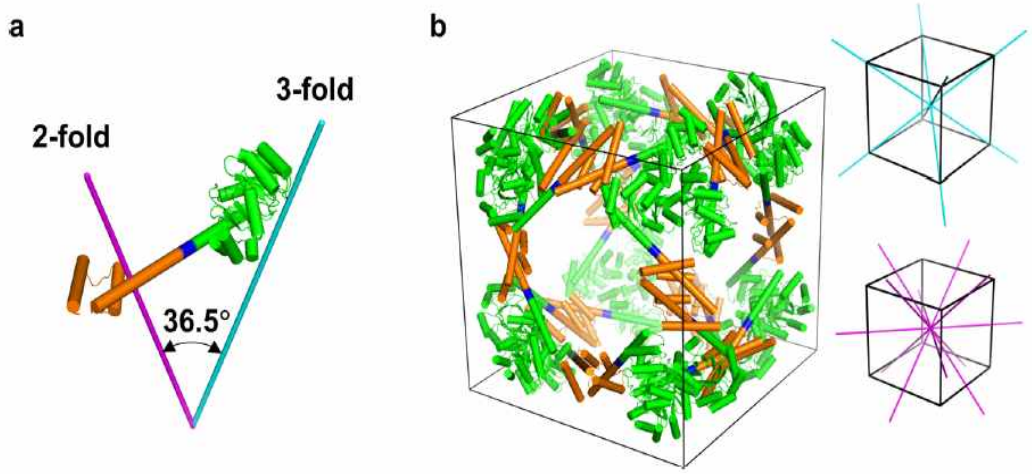
그중에 한 경우에 성공적인 케이지가 형성되는 것을 x-ray 크리스탈 구조를 통하여 확인할 수 있었다 (그림 2) (Lai et al., 2012a; Lai et al., 2012b; Lai et al., 2013). 이들이 디자인한 케이지는 각 꼭지점당 3개씩 12개의 BPO-M1 융합단백질로 형성된 케이지였다. 각 꼭지점에 있는 세 개의 헬릭스 링커 중 한 경우에는 완벽에 가까운 헬릭스 구조를 보여주었지만, 다른 두 경우에는 조금 꺾인 형태의 헬릭스 구조가 이루어졌다. 이 결과는 두 단백질을 연결하는 헬릭스의 구조적인 안정성과 예측 가능성이 쉽지 않다는 것으로 보여준다. 이 연구가 성공한 가장 큰 요인은 헬릭스 링커의 선택에 있었다. 연구팀은 오랫동안 비슷한 디자인을 시도하였지만 대부분의 경우 만들어진 단백질은 일정한 구조를 가지지 못하는 aggregate 가 된 단백질이었다. 하지만, 링커로 사용한 아미노산 중, 두 개의 아미노산이 새로운 아미노산으로 바뀌면서 링커의 헬릭스구조가 안정화되었고 결국 케이지 형성이 가능하였다고 한다.



<그림 2> 알파헬릭스를 이용한 tetrahedral 단백질 케이지 디자인. Trimer를 이루는 BPO (녹색) 단백질과 M1 (보라) 단백질을 알파헬릭스로 연결하여 융합단백질을 만들었다 (Lai et al., 2012a).

Yeates 그룹은 비슷한 방법으로 정육면체 형태의 단백질 케이지 형성에도 성공하였다 (Lai et al., 2014). 이들은 trimer를 형성하는 대장균의 2-keto-3-deoxy-6-phosphogalactonate (KDPGal) 단백질과 dimer를 형성하는 대장균 FkpA 단백질을 알파헬릭스 링커로 연결하였다. 연구자들은 5개의 다른 헬릭스 링커를 테스트하였으며, 그 중 한 경우에서 안정적인 구조를 가지는 단백질 복합체가

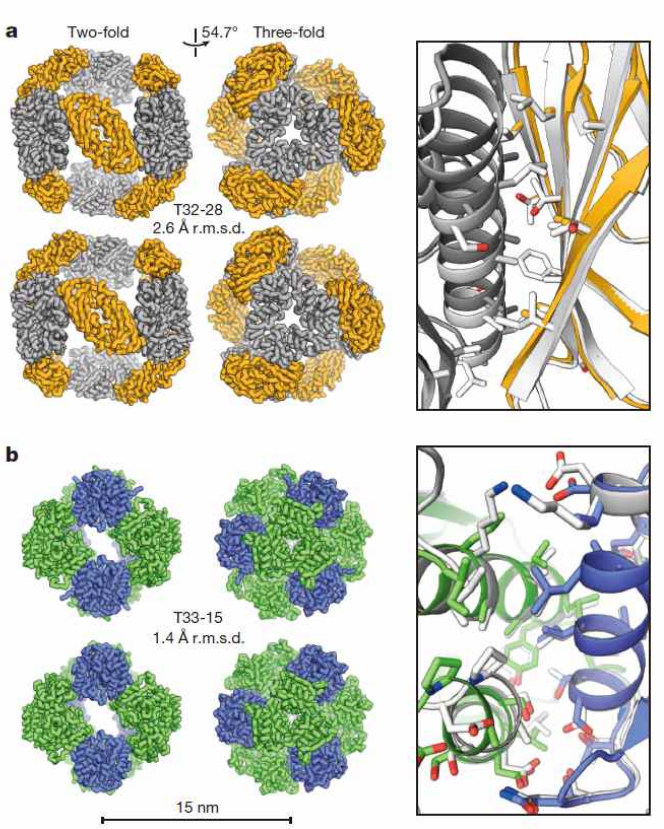
만들어지는 것을 확인하였다. 이 단백질의 x-ray 구조를 풀어본 결과 정확히 예측한 것과 같은 cube 형태의 다공성 구조를 형성한다는 것을 증명할 수 있었다.



<그림 3> 알파헬릭스 링커를 이용한 정육면체 단백질 케이지 디자인. Trimer를 이루는 KDPGal (녹색) 단백질과 dimer를 이루는 FkpA 단백질을 알파헬릭스 링커로 연결하였다 (Lai et al., 2014).

3. 단백질 인터페이스 디자인을 이용한 케이지 구조 형성

David Baker 연구팀은 앞에서 설명한 알파헬릭스 링커를 이용한 디자인과는 전혀 다른 방법으로 케이지 형태의 단백질 나노구조체 형성을 성공시켰다 (King et al., 2012). 이들은 Rosetta Design 프로그램을 이용하여, 자연적으로 trimer를 형성하는 많은 수의 단백질 구조를 Protein Data Base (PDB)에서 골라내고, 이것들이 정사면체 복합체를 형성할 수 있도록 docking을 시도하였다. 또한 Docking 계산을 수행하면서 각 단백질 사이의 interface 지역의 아미노산 서열을 최적화하면서 interface가 안정적으로 형성되도록 유도하였다. 연구팀은 이렇게 디자인된 복합체가 정확한 케이지 구조를 형성하는지 확인하기 위하여, 다섯 개의 단백질을 negative stain 전자현미경 이미지 분석을 하였으며, 이중 4개 단백질의 x-ray 크리스탈 구조를 규명하였다 (그림 4). 만들어진 단백질들은 모두 예상하였던 구조와 매우 유사한 구조를 형성하였으며, 이들의 디자인 방법이 유효하다는 것을 증명하였다.

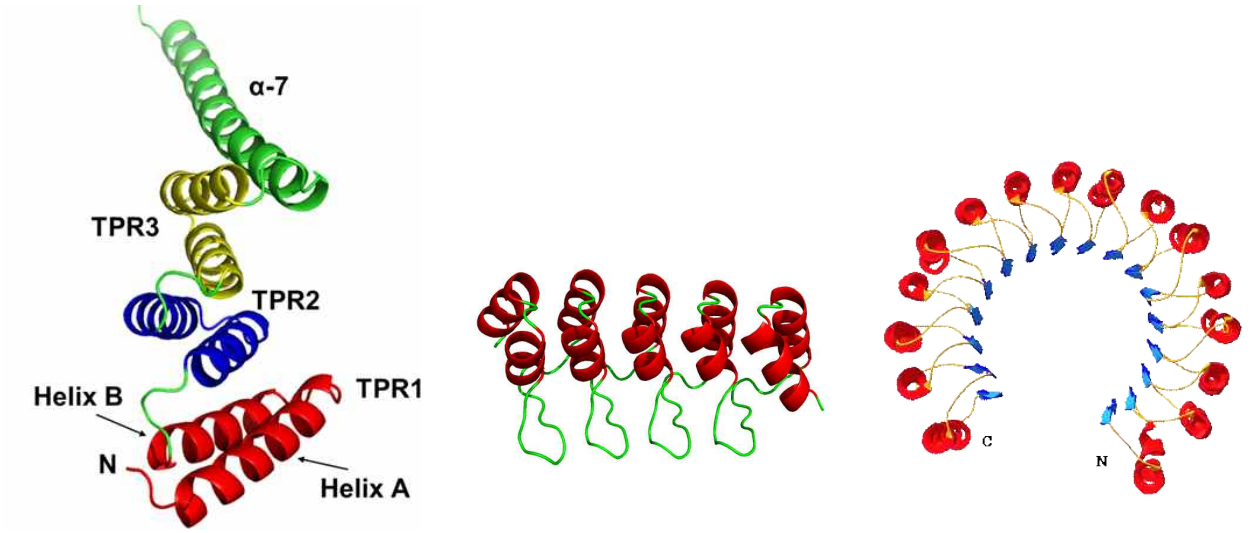


<그림 4> 인터페이스 최적화 방법으로 디자인한 단백질 케이지들의 x-ray 크리스탈 구조. 예상한 구조와 매우 유사한 단백질 케이지가 성공적으로 형성되었다 (King et al., 2012).

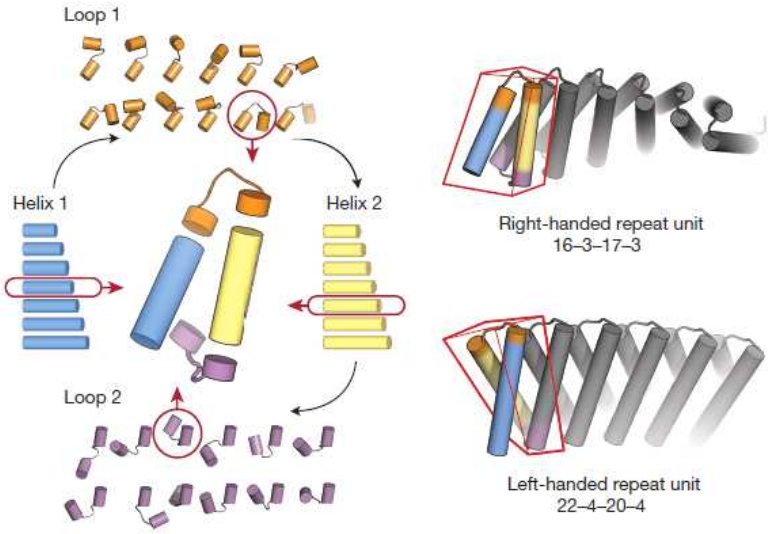
4. 인터페이스 디자인을 이용한 반복단백질 설계

반복단백질은 ankyrin repeat, Leucine rich repeat, Tetratricopeptide repeat 와 같은 20~30개의 짧은 펩티드 모티프가 계속 반복되어 있는 구조이다 (그림 5). 자연계에서 반복 단백질들은 다른 단백질과 결합하는 역할을 하는데 이러한 역할을 담당하는 이유는, 반복 단백질들이 반복 모티프의 구조를 안정시키는 부분과 기능을 담당하는 부분이 뚜렷하게 구분이 되기 때문에 비교적 쉽게 새로운 결합 기능의 진화가 가능하기 때문으로 보인다. 반복단백질은 세포에서 매우 다양한 역할을 하고 있으며, 비교적 쉽게 다른 단백질과 결합하는 바인더를 디자인할 수 있다는 장점이 있기 때문에 생명공학분야에서도 응용가능성이 높다. Baker와 Bradley 그룹은 인공 케이지 디자인에서와 유사하게 Rosetta Design 프로그램을 이용하여 반복단백질 구조디자인도 시도하였다 (Brunette et al., 2015; Doyle et al., 2015). 연구팀은 “알파헬릭스-루프-알파헬릭스”로 이루어진 반복모듈 모티프를 디자인하였는데 모듈사이의 인터페이스를 최적화함으로써 매우 안정적이면서, 자연계에는 존재하지 않는 새로운 반복단백질을 설계할 수

있었다 (그림 6). 이들은 83개의 반복단백질을 디자인하였는데, 그중에서 절반 정도인 44개 반복단백질에서 설계한 구조와 유사한 구조가 만들어졌다는 것을 확인하였다.



<그림 5> 자연에 존재하는 반복단백질. 왼쪽부터 TPR (tetratricopeptide repeat), ankyrin repeat, leucine rich repeat 단백질 구조.

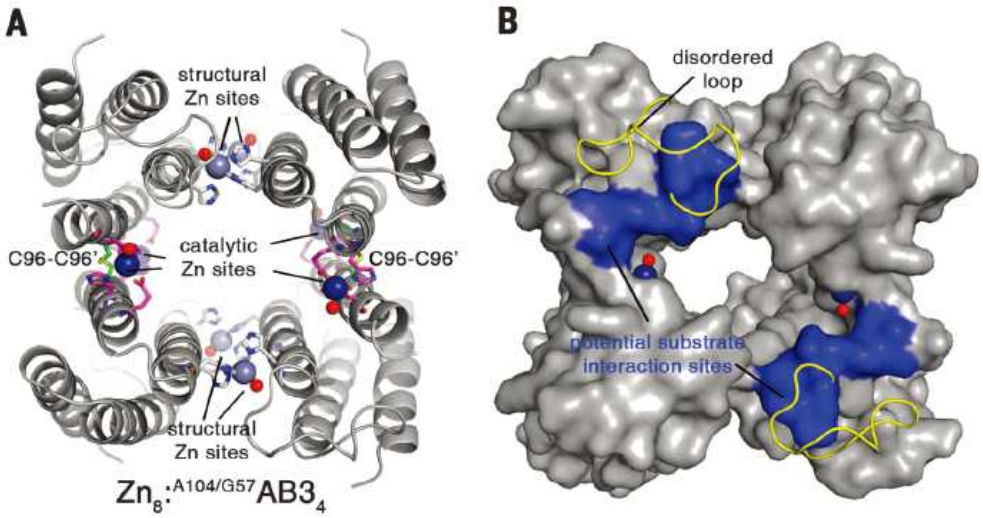


<그림 6> 인터페이스 설계방법을 이용한 알파헬릭스 반복단백질 디자인 방법 (Brunette et al., 2015).

5. 금속이온 결합구조를 이용한 단백질 복합체 디자인

단백질은 glutamate, aspartate, cysteine, histidine과 같이 배위결합이 가능한 아미노산을 이용하여 다양한 금속이온과 결합한다. 단백질에 결합된 금속이온은 효소의 촉매작용을 매개하기도 하며, 단백질의 구조적인 안정성에 기여하기도 한다. 금속이온의 배

위결합은 여러 개의 아미노산이 예측가능한 일정한 구조를 가져야만 형성이 되는데 이러한 금속이온 결합을 이용하면 두 개의 단백질에 있는 아미노산들을 일정한 구조로 배열할 수 있다. 이러한 성질을 이용하여 Akif Tezcan 그룹은 다양한 단백질 구조체를 디자인하였다 (Radford et al., 2010a; Radford et al., 2010b; Radford and Tezcan, 2009). 연구팀은 노출된 알파헬릭스 한쪽 면에 histidine이나 cysteine 그룹을 배치하여 단백질의 아미노산만으로 zinc나 copper 결합부위를 디자인하였다. Zinc나 copper는 네 개의 아미노산과 결합하는데, 하나의 단백질에 두 개의 아미노산을 배치하면, 비어있는 2개의 배위결합자리를 다른 단백질이 채워주면서 단백질의 결합을 유도할 수 있게 된다. 연구팀은 이러한 설계방법을 응용하여 인공적인 단백질 활성부위를 설계할 수 있었으며, 여기에 새로운 촉매기능을 구현할 수도 있었다 (그림 7) (Song and Tezcan, 2014). 연구팀은 또한 자연적인 단백질에는 존재하지 않는 인공화합물인 Phen과 Quinn을 단백질의 cysteine에 conjugation 시켜서 인공적인 금속결합부위를 설계하였다 (그림 8). 이러한 방법을 사용하여 단백질의 원하는 부위에 선택적으로 금속결합부위를 배치할 수 있었고, 예측가능한 단백질 복합체를 설계할 수 있었다 assemblies (Brodin et al., 2012; Brodin et al., 2010; Medina-Morales et al., 2013; Salgado et al., 2010; Salgado et al., 2007; Salgado et al., 2008; Song et al., 2014). 하지만 금속이온을 이용한 방법은 서로 다른 단백질 heterodimer 형성을 유도하는 것이 쉽지 않고, 두 단백질의 결합방향을 조절하는 것이 쉽지 않다는 단점이 있다. 그렇지만, metal chelation 방법은 metal을 이용하여 단백질 복합체의 구조를 조절할 수 있다는 큰 장점이 있기 때문에 외부에서 조절이 가능한 구조체를 설계할 수 있는데 이러한 성질을 이용하여 Tezcan 그룹은 금속이온의 농도를 변화시켜서 단백질 케이지 형성을 on-off 조절할 수 있는 페리틴을 설계할 수 있었다.



<그림 7> Zinc 결합부위를 이용한 인공효소 설계. Zinc 결합부위를 이용하여 단백질 구조를 만들고, 여기에 활성부위를 설계하였다 (Song and Tezcan, 2014).

<그림 8> Phen 그룹을 이용한 금속결합 부위 설계. 금속이온 결합이 두 단백질 헬릭스의 특이적인 결합을 유도한다 (Radford et al., 2010b).

6. 고찰

분자의 자가조립 특성을 이용하여 다양한 모양과 기능을 가지는 나노구조체를 만드는 연구는 지난 30년간 많은 발전을 하였다. 유기분자와 금속을 이용한 MOF (metal organic framework) 구조를 만들 수 있으며, DNA의 자가조립 성질을 이용하면 더욱 다양한 구조를 만들어낼 수 있다. 단백질은 대부분의 세포기능을 담당하고 있는 분자이며, 구조적으로 화학적 특성 면에서 다른 분자와는 비교할 수 없는 다양성을 가지고 있다. 이러한 구조적인 다양성과 비교적 약한 화학결합으로 구조가 유지된다는 점 때문에 그동안 인공단백질을 이용한 새로운 나노구조체 연구는 큰 진전을 이루기가 쉽지 않았다. 하지만 최근, 특히, 알파헬릭스가 가지는 구조적인 안정성과 독립성을 이용하여 안정적이고 예측가능한 단백질 구조체 디자인이 가능하다는 것이 속속 밝혀지고 있다. 단백질을 이용한 나노구조체 연구는 DNA나 다른 생체분자 나노구조체에 비하여 대량 생산이 용이하고, 보다 다양한 기능을 구현할 수 있기 때문에 단백질치료제, 항체, 인공효소, 백신디자인 등 다양한 분야에서 응용이 가능할 것으로 기대된다.

참고문헌

1. Brodin, J.D., Ambroggio, X.I., Tang, C., Parent, K.N., Baker, T.S., and Tezcan, F.A. (2012). Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. Nature chemistry *4*, 375-382.

2. Brodin, J.D., Medina-Morales, A., Ni, T., Salgado, E.N., Ambroggio, X.I., and Tezcan, F.A. (2010). Evolution of metal selectivity in templated protein interfaces. J Am Chem Soc *132*, 8610-8617.

3. Brunette, T.J., Parmeggiani, F., Huang, P.S., Bhabha, G., Ekiert, D.C., Tsutakawa, S.E., Hura, G.L., Tainer, J.A., and Baker, D. (2015). Exploring the repeat protein universe through computational protein design. Nature *528*, 580-584.

4. Doyle, L., Hallinan, J., Bolduc, J., Parmeggiani, F., Baker, D., Stoddard, B.L., and Bradley, P. (2015). Rational design of alpha-helical tandem repeat proteins with closed architectures. Nature *528*, 585-588.

5. King, N.P., Sheffler, W., Sawaya, M.R., Vollmar, B.S., Sumida, J.P., Andre, I., Gonen, T., Yeates, T.O., and Baker, D. (2012). Computational design of self-assembling protein nanomaterials with atomic level accuracy. Science *336*, 1171-1174.

6. Lai, Y.T., Cascio, D., and Yeates, T.O. (2012a). Structure of a 16-nm cage designed by using protein oligomers. Science *336*, 1129.

7. Lai, Y.T., Jiang, L., Chen, W., and Yeates, T.O. (2015). On the predictability of the orientation of protein domains joined by a spanning alpha-helical linker. Protein Eng Des Sel *28*, 491-500.

8. Lai, Y.T., King, N.P., and Yeates, T.O. (2012b). Principles for designing ordered protein assemblies. Trends Cell Biol *22*, 653-661.

9. Lai, Y.T., Reading, E., Hura, G.L., Tsai, K.L., Laganowsky, A., Asturias, F.J., Tainer, J.A., Robinson, C.V., and Yeates, T.O. (2014). Structure of a designed protein cage that self-assembles into a highly porous cube. Nature chemistry *6*, 1065-1071.

10. Lai, Y.T., Tsai, K.L., Sawaya, M.R., Asturias, F.J., and Yeates, T.O. (2013). Structure and flexibility of nanoscale protein cages designed by symmetric self-assembly. J Am Chem Soc *135*, 7738-7743.

11. Medina-Morales, A., Perez, A., Brodin, J.D., and Tezcan, F.A. (2013). In vitro and cellular self-assembly of a Zn-binding protein cryptand via templated disulfide bonds. J Am Chem Soc *135*, 12013-12022.

12. Padilla, J.E., Colovos, C., and Yeates, T.O. (2001). Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proceedings of the National Academy of Sciences of the United States of America *98*, 2217-2221.

13. Radford, R.J., Nguyen, P.C., Ditri, T.B., Figueroa, J.S., and Tezcan, F.A. (2010a). Controlled protein dimerization through hybrid coordination motifs. Inorg Chem *49*, 4362-4369.

14. Radford, R.J., Nguyen, P.C., and Tezcan, F.A. (2010b). Modular and versatile hybrid coordination motifs on alpha-helical protein surfaces. Inorg Chem *49*, 7106-7115.

15. Radford, R.J., and Tezcan, F.A. (2009). A superprotein triangle driven by nickel(II) coordination: exploiting non-natural metal ligands in protein self-assembly. J Am Chem Soc *131*, 9136-9137.

16. Salgado, E.N., Ambroggio, X.I., Brodin, J.D., Lewis, R.A., Kuhlman, B., and Tezcan, F.A. (2010). Metal templated design of protein interfaces. Proceedings of the National Academy of Sciences of the United States of America *107*, 1827-1832.

17. Salgado, E.N., Faraone-Mennella, J., and Tezcan, F.A. (2007). Controlling protein-protein interactions through metal coordination: assembly of a

16-helix bundle protein. J Am Chem Soc *129*, 13374-13375.

18. Salgado, E.N., Lewis, R.A., Faraone-Mennella, J., and Tezcan, F.A. (2008). Metal-mediated self-assembly of protein superstructures: influence of secondary interactions on protein oligomerization and aggregation. J Am Chem Soc *130*, 6082-6084.

19. Song, W.J., Sontz, P.A., Ambroggio, X.I., and Tezcan, F.A. (2014). Metals in protein-protein interfaces. Annual review of biophysics *43*, 409-431.

20. Song, W.J., and Tezcan, F.A. (2014). A designed supramolecular protein assembly with in vivo enzymatic activity. Science *346*, 1525-1528.