

컴퓨터를 이용한 생체 고분자 상 분리 모사

Computer Simulations of Biomacromolecular Phase Separation

엄유곤 · 최정모 | Yu-Gon Eom · Jeong-Mo Choi

Department of Chemistry, Pusan National University,
2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241, Korea
E-mail: jmchoi@pusan.ac.kr

1. 서론

상 분리(phase separation)는 용액 내 성분들이 자발적으로 분리되어 여러 개의 상을 형성하는 열역학적 과정이다. 생체 고분자 상 분리에서 시스템은 단백질과 같은 생체 고분자와 물(또는 완충액)의 혼합물로 이해할 수 있으며, 많은 경우 생체 고분자가 적은 상과 생체 고분자가 풍부한 상으로 분리된다.^{1,2} 상 분리 현상은 엔트로피와 엔탈피 사이의 경쟁에 의해 결정된다. 일반적인 이성분 계에서 엔트로피는 분자들의 혼합을 선호하는 반면, 엔탈피는 특정 분자 상호작용의 형성을 촉진한다. 따라서, 엔탈피적 이득이 엔트로피적 페널티보다 우세할 때, 계는 자발적으로 상 분리를 겪게 된다.

세포는 단백질 및 다른 생체 고분자의 상 분리를 이용하여 다양한 작업을 수행할 수 있다. 특정한 분자들이 공간에서 모이거나 배제될 수 있으며, 이를 통해 세포 내 반응 속도를 조절할 수 있다. 각 상 내의 농도가 일정하게 유지된다는 상 분리의 특성을 이용하여 완충 작용을 할 수도 있고, 표면 장력으로 힘을 만들어 내기도 한다.

세포 내 상 분리를 추동하는 생체 고분자들 중에서 단일 분자만으로도 상 분리를 일으킬 수 있는 대표적인 분자로는 멀티도메인 단백질(multi-domain protein)과 비정형 단백질(intrinsically disordered protein)이 있다. 멀티도메인 단백질은 여러 단백질 도메인이 비정형 링커로 연결된 구조를 가지고 있으며, 비정형 단백질은 안정적인 3차원 구조 없이 다양한 구조를 가질 수 있다. 이러한 분자들은 스티커 및 스페이서 모형(stickers-and-spacers model)을 통해 이해할 수 있다(그림 1).³ 스티커는 다른 고분자와의 상호작용을 담당하는 영역이고, 스페이서는 상대적으로 상호작용에 참여하지 않는 영역이다. 생체 고분자의 상 분리에서는 이러한 고분자들의 약한 다가성(multivalent) 상호작용이 상 분리의 추진력으로 작용한다고 알려져 있다. 본 기고문에서는 이러한 상 분리 과정을 컴퓨터를 활용하여 모사하는 여러 접근법을 소개하고, 그 응용 사례를 논하고자 한다.

2. 본론

2.1 입자 기반 시뮬레이션

입자 기반 시뮬레이션은 분자를 입자들로 묘사하고 물리 법칙을 적용하여 계의 움직임을 표현한다. 입자 간의 상호작용은 관련 입자의 유형과 공간상 위치에 따라 달라지며, 수학적으로 정의될 수 있다. 계를 열적 평형에 이르게 하거나 동력학적 거동을 모델링하기 위한 방법으로는 랜덤 샘플링을 사용한 확률 계산을 통해 계의

Author



엄유곤

2022

2024

2024-현재

부산대학교 바이오소재학과 (학사)
부산대학교 화학과 (석사)
부산대학교 화학과 (박사과정)



최정모

2011

2015

2016

2020-현재

KAIST 화학과 (학사)
Harvard University, History of Science (석사)
Harvard University, Chemistry and Chemical Biology (박사)
부산대학교 화학과 조교수

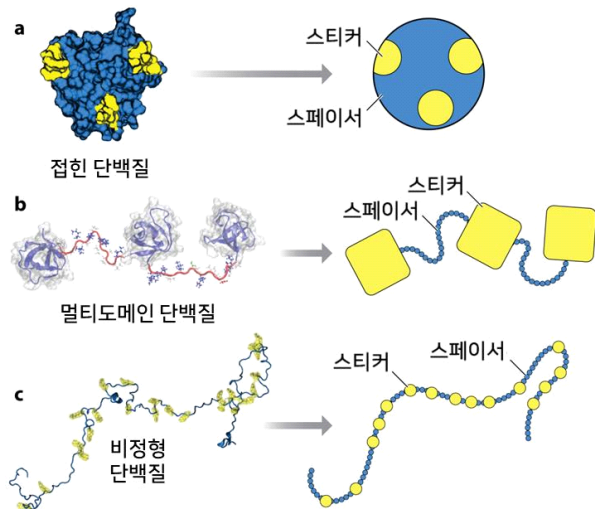


그림 1. 다양한 단백질 시스템의 스티커와 스페이서 유형 개요.³ (a) 접힌 단백질, (b) 멀티도메인 단백질, (c) 비정형 단백질.

변화를 추적하는 몬테 카를로(Monte Carlo) 방식과 뉴턴의 운동 방정식을 적분해 입자의 위치, 속도, 가속도를 시간의 함수로 표현하는 분자 동역학(molecular dynamics) 방식이 있다.

입자 기반 시뮬레이션은 묘사 수준에 따라 원자 수준 시뮬레이션과 coarse-grained 시뮬레이션으로 분류할 수 있다. 원자 수준 시뮬레이션은 원자 수준에서 분자 계를 모델링하며 매우 상세한 정보를 제공한다(그림 2a). 하지만 상당한 계산 자원을 요구하며, 특히 상 분리 같은 집단 거동은 적절한 모델링을 위해 많은 수의 분자를 포함해야 하기 때문에 현재의 계산 방법론 내에서 시뮬레이션을 실행하는 것에는 많은 어려움이 따른다. 따라서 상 분리 연구에서 원자 수준 시뮬레이션의 사용은 적은 수의 짧은 생체 고분자들에 한정된다. Coarse-grained 시뮬레이션은 단백질의 아미노산 잔기와 같이 개별 원자나 분자를 더 적은 수, 더 큰 단위로 대체하는 컴퓨터 모델링 전략으로, 계산 비용을 줄이는데 그 목표가 있다(그림 2b). 상 분리 모델링 coarse-grained 시뮬레이션은

공간적 표현 방식에 따라 격자 모델과 비격자 모델로 나눌 수 있다.

2.1.1 격자 모델

격자 모델은 2차원 또는 3차원 격자 공간을 사용해서 계산 비용을 줄이는 것이 특징이다(그림 2c). 입자들은 오직 격자점에만 위치할 수 있으며, 이동 시에는 한 격자점에서 다른 격자점으로의 이동만 허용된다. 합성 고분자의 거동을 연구하기 위해 많은 격자 모델이 개발되어 왔으며, 그 중 일부는 생체 고분자 상 분리 연구에도 사용되고 있다.

일부 격자 모델은 주로 멀티도메인 단백질 모델링을 목표로 하는데, 여기서 접힌 도메인은 격자점에 위치한 입자로 모델링된다.⁴ 이 때, 비정형 링커는 암시적으로(implicitly) 또는 명시적으로(explicitly) 모델링될 수 있다. 암시적 모델은 링커에게 격자점을 할당하지 않으며, 링커는 입자 간 거리를 제한하는 역할만 한다. 반면, 명시적 모델에서는 링커가 격자점을 차지하는 비활성 입자로 표현된다. 격자 모델의 응용 사례로, 멀티도메인 단백질의 단순한 모형을 가정하여 링커의 성질이 상 분리에서 어떤 역할을 수행하는지 살펴본 연구가 있으며, 유사한 모델이 NPM1-FIB1 응축물의 층상 구조를 재현하고 설명하기 위해 적용된 바 있다.⁵ 또다른 격자 기반 시뮬레이션은 고분자의 유연성에 따른 상 거동을 예측하는데 사용되었고, 고분자의 유연성이 상 분리 시스템에서 “마법수” 효과를 강화한다는 것을 발견했다.⁶

보다 보편적으로 이용할 수 있는 격자 기반 생체 고분자 상 분리 시뮬레이션 엔진도 개발되고 있다. PIMMS(Polymer Interactions in Multi-components Mixtures)는 비정형 단백질의 상 분리를 체계적으로 모델링할 수 있는 격자 기반 시뮬레이션 엔진이다.⁷ PIMMS는 hnRNP A1 단백질 프리온 유사 도메인의 상 거동을 방향족 잔기의 수와 패턴에 따라 살펴보는 데 이용되었으며(그림 3),⁸ Velo1 단백질의 상 분리도 모델링하였다.⁹ 또 다른 격자 시뮬레이션 엔진으로는, 스티커 및 스페이서 모형에 기반을 둔 LASSI(LATtice simulation

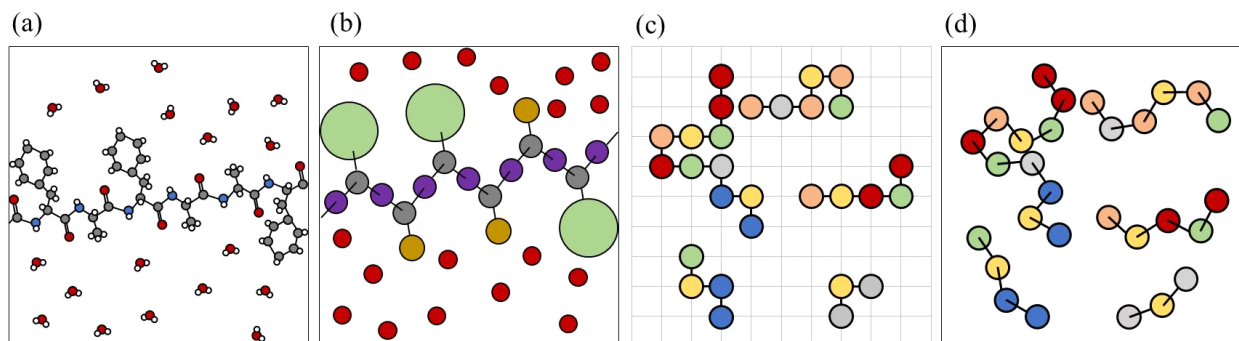


그림 2. 다양한 시뮬레이션 방식을 묘사한 그림. (a) 원자 수준 시뮬레이션, (b) coarse-grained 시뮬레이션, (c) 격자 모델, (d) 비격자 모델. 다양한 색상은 다른 유형의 입자를 나타낸다. 검은 실선은 공유 결합을 나타내며, 회색 실선은 격자선을 나타낸다.

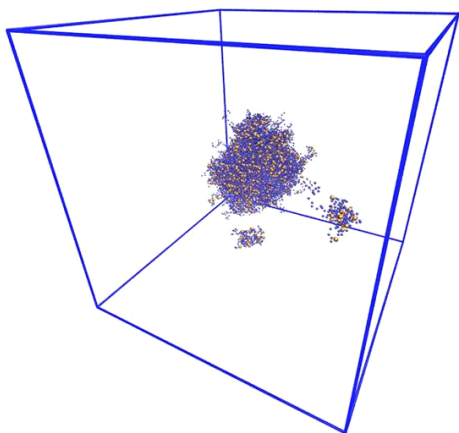


그림 3. 격자 모델인 PIMMS로 수행한 상 분리 시뮬레이션의 한 장면.⁷ 퇴행성 뇌질환에 관여하는 것으로 알려진 단백질 hnRNP A1의 상 분리 과정을 묘사하였다.

engine for Sticker and Spacer Interactions)가 있다.⁴ LASSI를 통해서 멀티도메인 단백질뿐 아니라 비정형 단백질, RNA 분자의 상 분리도 모델링할 수 있다. LASSI는 단백질 시스템, 단백질-RNA 혼합 시스템의 상 분리 연구에 널리 활용되었다.¹⁰⁻¹² 구체적인 응용 사례로는, 미토콘드리아 RNA가 생체 응축체에 미치는 영향을 연구하기 위해 적용된 바 있다.¹³

2.1.2 비격자 모델

비격자 모델은 입자들을 3차원 실제 공간에 배치하여 입자의 복잡한 운동을 묘사할 수 있다는 점에서 더 현실적이라고 할 수 있다(그림 2d). GROMACS, LAMMPS, HOOMD-blue 같은 시뮬레이션 엔진들은 비격자 시뮬레이션에 널리 사용되며, coarse-grained 시뮬레이션뿐 아니라 원자 수준 시뮬레이션에도 좋은 프레임워크로 작용한다. 이러한 엔진들을 활용하여 생체 고분자의 상 분리 현상을 분석하기 위한 다양한 모델이 개발되었다.

Coarse-grained 시뮬레이션의 경우, 원자들을 어떻게 그룹화하여 더 큰 입자로 묘사하는지에 따라 MARTINI, SIRAH, Kim-Hummer(KH) 모델, hydrophobicity scale(HPS) 모델, modified liquid droplet(MLD) 모델 등 다양한 coarse-grained 모델이 개발되어 있다. 다만 상호작용 매개변수를 결정하는 것이 단순한 일이 아니기에, 각 모델은 자체적인 가정에 의존하게 된다. 최근 기계 학습 기반 coarse-graining 접근법으로 CAMELOT이 개발되었는데, 이 CAMELOT은 coarse-grained 모델의 매개변수를 최적화하는데 기계학습을 사용하여 인간의 개입을 최소화하였으며, 이성분계 상 분리 모델링에 적용되어 응축물의 구조적 특성을 설명하는 데 사용된 바 있다.¹⁴

Coarse-grained 시뮬레이션은 입자의 규모에 따라 분류될 수 있다. 가장 작은 규모로 알려진 모델은 각 아미노산 잔기를 뼈대(backbone)와 곁사슬(side chain) 두 가지로 분류하여

두 입자가 한 잔기를 표현하는 모델이다. 그 중 하나로, 개량된 MARTINI 2.2 역장을 사용해 FUS의 프리온 유사 도메인의 상 분리를 묘사하는 시뮬레이션이 진행된 바 있다.¹⁵ 유사하게, MARTINI 3 역장을 사용하여 단백질 올리고머의 상 거동에 대한 pH 및 이온 농도의 영향을 연구한 사례가 있다.¹⁶

각 잔기에 하나의 입자를 할당하는 방식으로 모델링이 이루어지기도 한다. 예를 들어 MLD모델을 사용하여 FUS-LC와 NDDX4의 물리화학적 성질을 비교한 연구가 있으며, 이를 통해 FUS-LC가 NDDX4보다 더 높은 상 분리 경향성을 가진다는 것이 입증되었다.¹⁷ FUS와 LAF-1의 서열 변화를 HPS모델과 KH모델로 모델링하여 상 도표의 변화를 분석하기도 하였다.¹⁸ FUS와 RNA의 시뮬레이션에도 활용되어, RNA가 다가성 결합 분자로서 단백질과 상호작용하여 단백질의 상 분리를 촉진한다는 것이 밝혀졌다.¹⁹

입자 규모가 더 큰 시뮬레이션도 있으며, 이 경우 입자 하나는 두 개 이상의 잔기를 묘사한다. 일례로, 단백질과 RNA의 상호작용을 모델링하여 전사 후 수정(posttranslational modification)과 같은 변화가 심각한 액체-고체 전이를 유발할 수 있다는 것이 밝힌 연구가 있다.²⁰ 또 다른 시뮬레이션으로, 두 가지 다른 유형의 동종중합체로 이루어진 혼합물에서 화학양론에 따른 상 분리 효과에 대한 원리를 연구한 바 있다.²¹ 또한, 서열에 따라 비정형 단백질의 상 거동이 어떻게 변화하는지 시뮬레이션한 사례도 있다.²²

2.2 장 이론적 방법

장 이론적 시뮬레이션(field-theoretic simulation, FTS)은 고분자를 밀도 장(density field)으로 기술해 계의 열역학적 성질을 계산하는 방법이다.²³ Complex Langevin 샘플링을 이용한 FTS(FTS-CL)는 단일 고분자를 다른 고분자가 만들어내는 “장”에서 고려함으로써 계산적으로 어려운 다체 상호작용을 피한다. FTS-CL은 입자 기반 모델에서 시작해, 입자 기반 정준 분배 함수를 장 이론 분배 함수로 변환한다. 이렇게 단일 사슬에 대한 분배 함수로 나타나는 계의 해밀토니안은 Langevin동역학을 사용하여 수치적으로 샘플링된다. FTS-CL은 화학 퍼텐셜과 고분자 농도에 대한 정보를 제공함으로써 상 분리 상 도표를 직접적으로 얻을 수 있게 해준다.

FTS는 근사가 필요 없는 샘플링을 가능하게 하여 그 오차를 크게 줄일 수 있다. 입자 기반 시뮬레이션과 비교할 때, FTS는 평형에 빠르게 도달하므로 효율적인 계산을 가능하게 한다. 그러나 간단한 격자 단백질 시뮬레이션과 그에 상응하는 FTS-CL을 비교했을 때, 배제 부피(exclusive volume)의 반발력이 너무 강하지 않은 경우에 한해서만 결과가 재현 가능한 수준으로 나타났다.²⁴

FTS는 tau-RNA 응축물의 상 도표를 얻는 데 사용된 바

있다. 생리적 조건에서 tau-RNA의 상 경계 위치가 상 거동의 열역학적 원리를 설명하며, 이로써 세포 내의 작은 변화조차 상 분리나 단백질 응집을 유도할 수 있다는 것을 보여주었다.²⁵ 또 다른 적용 사례에서는, FTS 계산으로 얻은 상 도표가 미생물의 세포 과립 단백질 Ddx4에 대한 실험 결과와 일치하는 것으로 나타났다.^{26,27}

2.3 기계 학습 기반 방법

인공 지능 분야에서 딥러닝 기법이 등장한 이후, 기계 학습은 보다 정확한 분류 및 예측을 할 수 있는 능력으로 인해 많은 관심을 받아왔다. 다양한 기계 학습 접근법이 과학 연구에 적용되고 있으며, 잘 알려진 예로는 단백질 구조 예측 모델 AlphaFold와 단백질 구조 생성 모델 RFDiffusion이 있다. 같은 맥락에서, 최근 생체 고분자의 상 분리 패턴을 예측하기 위해 기계 학습 방법에 대한 여러 시도가 이루어졌다. “모델”에 의존하지 않는 기계 학습 기반 접근 방식은 전통적인 모델 기반 접근 방식보다, 임의의 분자에 대한 상 분리 경향성 예측과 같은 특정 작업을 훨씬 빠르게 수행할 수 있다는 것이 장점으로 꼽힌다.

기계 학습 알고리즘은 데이터로부터 학습을 하며, 이에 기반하여 데이터를 분류하거나 예측값을 제공한다. 따라서 생체 고분자 상 분리의 기계 학습 예측을 위해서는 기계가 학습할 데이터 세트가 필요하며, 다음 데이터베이스들이 자주 사용된다. PhaSepDB는 기관체에서 상 분리와 관련된 단백질을 포함하는 데이터베이스로, 출판된 문헌과 다른 데이터베이스에서 수집한 자료로 구성되었다. 마찬가지로, LLPsDB는 상 분리에 관여하는 단백질과 그들의 상 거동 및 실험 조건에 대한 정보를 제공하는 데이터베이스다. PhaSePro는 실험적으로 검증된 상 분리 발현 단백질과 단백질 도메인을 포함하는 수동 큐레이션 데이터베이스다. DisProt은 실험적으로 결정된 비정형 도메인을 가진 단백질에 대한 정보를 포함하는 데이터베이스다. UniProt과 Protein Data Bank(PDB)와 같은 고전적인 서열 및 구조 데이터베이스도 학습 과정에서 자주 활용된다.

위 데이터베이스를 토대로, 임의의 단백질 서열의 상 분리 경향을 예측하기 위한 여러 지도 학습 알고리즘이 개발되었다. 그 중 하나로, UniProt과 PhaSePro의 데이터셋을 통해 학습된 로지스틱 회귀 모델 FuzDrop²⁸이 있다. DeePhase²⁹와 PSAP³⁰은 LLPsDB와 PhaSepDB의 데이터셋으로 훈련된 랜덤 포레스트 분류 엔진이며, DeePhase²⁹에서는 단백질 서열 처리를 위해 word2vec 기술이 사용되었다. 마찬가지로, 그래디언트 부스팅 결정 트리 모델인 PSPredictor³¹에서도 word2vec이 사용되었으며 데이터베이스로는 LLPsDB가 사용되었다. PScore³²는 PDB의 데이터셋으로 훈련된 선형 회귀 모델로, 상 분리 경향 계산에 π - π 상호작용을 포함한다. 이는 나중에

삼중 모델인 LLPhyScore³³로 개선되었다. Paiz 등은 PhaSePro와 DisProt의 데이터셋으로 ParSe³⁴를 개발했고, 그 알고리즘은 베타 구조 경향성의 계산을 상 분리 경향 예측에 사용한다. Ibrahim 등은 알파 구조 경향성 계산을 포함시켜 알고리즘을 개선했다.³⁵

데이터베이스와 기계 학습 기법의 발전에도 불구하고, 데이터의 편향성은 해결해야 할 숙제로 남아있다. 편향된 데이터로 훈련된 기계 학습 모델은 상당히 치우친 예측 결과를 제공할 수 있는데, 현재 생체 고분자 상 분리에 대한 데이터베이스는 FUS와 같이 소수의 잘 알려진 모델 시스템에서 대부분의 실험 데이터를 수집했다는 약점을 안고 있다. 또한, 전통적인 서열 및 구조 데이터베이스와 비교할 때, 상 분리 관련 데이터베이스의 크기는 상대적으로 작은 것이 현실이다.

3. 결론

세포 내 상 분리가 수행하는 역할에 대한 인식이 증가함에 따라, 이 분야에 대한 관심도 높아졌다. 계산적 접근법은 실험적 접근법을 보완하여 공간 및 시간에서의 ‘고해상도’ 데이터를 제공하며, 시뮬레이션 모델의 구성 요소를 임의로 변경함으로써 분자적 원리를 체계적으로 분석할 수 있게 해준다. 본 총설에서는 이들의 대표적인 예로서 입자 기반 시뮬레이션, 장 이론적 방법, 그리고 기계 학습 기반 방법에 대하여 소개하였다.

그러나 특히 살아 있는 세포 내에서의 상 분리를 모델링하는 측면에서 여전히 많은 질문이 남아 있다. 물리적, 화학적으로 끊임없는 변화를 겪는 복잡한 세포 내 환경을 시뮬레이션할 수 있는가? 세포 내에서도 같이 불균일하고 비평형인 조건에서 생체 고분자의 상 분리 경향을 예측할 수 있는가? 세포 내에서 상 분리와 능동적/수동적 세포 과정 사이의 상호작용도 모델링할 수 있는가? 우리는 이러한 질문에 답하기 위해 새로운 모델과 방법을 개발하고 적용하는 노력을 멈추지 않을 것이다.

참고문헌

1. C. P. Brangwynne, P. Tompa, and R. V. Pappu, *Nat. Phys.*, **11**, 899 (2015).
2. R. J. Emenecker, A. S. Holehouse, and L. C. Strader, *Annu. Rev. Plant Biol.*, **72**, 17 (2021).
3. J.-M. Choi, A. S. Holehouse, and R. V. Pappu, *Annu. Rev. Biophys.*, **49**, 107 (2020).
4. J.-M. Choi, F. Dar, and R. V. Pappu, *PLoS Comp. Biol.*, **15**, e1007028 (2019).
5. M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, and C. P. Brangwynne, *Cell*, **165**, 1686 (2016).

6. B. Xu, G. He, B. G. Weiner, P. Ronceray, Y. Meir, M. C. Jonikas, and N.S. Wingreen, *Nat. Commun.*, **11**, 1561 (2020).
7. A. S. Holehouse and R. V. Pappu, *Zenodo*, DOI: 10.5281/zenodo.3588456 (2019).
8. E. W. Martin, A. S. Holehouse, I. Peran, M. Farag, J. J. Incicco, A. Bremer, C. R. Grace, A. Soranno, R. V. Pappu, and T. Mittag, *Science*, **367**, 694 (2020).
9. A. S. Holehouse, G. M. Ginell, D. Griffith, and E. Böke, *Biochemistry*, **60**, 3566 (2021).
10. I. Seim, C. A. Roden, and A. S. Gladfelter, *Biophys. J.*, **120**, 2771 (2021).
11. M. Kar, F. Dar, T. J. Welsh, L. T. Vogel, R. Kühnemuth, A. Majumdar, G. Krainer, T. M. Franzmann, S. Alberti, C. A. M. Seidel, T. P. J. Knowles, A. A. Hyman, and R. V. Pappu, *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2202222119 (2022).
12. I. Seim, A. E. Posey, W. T. Snead, B. M. Stormo, D. Klotsa, R. V. Pappu, and A. S. Gladfelter, *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2120799119 (2022).
13. M. Feric, A. Sarfallah, F. Dar, D. Temiakov, R. V. Pappu, and T. Misteli, *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2207303119 (2022).
14. F. Dar, S. R. Cohen, D. M. Mitrea, A. H. Phillips, G. Nagy, W. C. Leite, C. B. Stanley, J.-M. Choi, R. W. Kriwacki, and R. V. Pappu, *bioRxiv*, DOI: 10.1101/2023.10.07.561338 (2023).
15. Z. Benayad, S. von Bülow, L. S. Stelzl, and G. Hummer, *J. Chem. Theory Comput.*, **17**, 525 (2021).
16. Y. Liu, X. Wang, Z. Wan, T. Ngai, and Y.-L. S. Tse, *Chem. Sci.*, **14**, 1168 (2023).
17. L. Li, M. Paloni, A. R. Finney, A. Barducci, and M. Salvalaglio, *J. Phys. Chem. Lett.*, **14**, 1748 (2023).
18. G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, *PLoS Comp. Biol.*, **14**, e1005941 (2018).
19. J. A. Joseph, J. R. Espinosa, I. Sanchez-Burgos, A. Garaizar, D. Frenkel, and R. Collepardo-Guevara, *Biophys. J.*, **120**, 1219 (2021).
20. S. Ranganathan and E. Shakhnovich, *Biophys. J.*, **121**, 2751 (2022).
21. A. G. T. Pyo, Y. Zhang, and N. S. Wingreen, *iScience*, **25**, 103852 (2022).
22. A. Statt, H. Casademunt, C. P. Brangwynne, and A. Z. Panagiotopoulos, *J. Chem. Phys.*, **152**, 075101 (2020).
23. G. H. Fredrickson, V. Ganesan, and F. Drolet, *Macromolecules*, **35**, 16 (2002).
24. D. Nilsson, B. Bozorg, S. Mohanty, B. Söderberg, and A. Irbäck, *J. Chem. Phys.*, **156** (2022).
25. Y. Lin, J. McCarty, J. N. Rauch, K. T. Delaney, K. S. Kosik, G. H. Fredrickson, J.-E. Shea, and S. Han, *eLife*, **8**, e42571 (2019).
26. J. P. Brady, P. J. Farber, A. Sekhar, Y.-H. Lin, R. Huang, A. Bah, T. J. Nott, H. S. Chan, A. J. Baldwin, J. D. Forman-Kay, and L. E. Kay, *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E8194 (2017).
27. K. T. Delaney and G. H. Fredrickson, *J. Chem. Phys.*, **146** (2017).
28. M. Hardenberg, A. Horvath, V. Ambrus, M. Fuxreiter, and M. Vendruscolo, *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 33254 (2020).
29. K. L. Saar, A. S. Morgunov, R. Qi, W. E. Arter, G. Krainer, A. A. Lee, and T. P. J. Knowles, *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2019053118 (2021).
30. G. van Mierlo, J. R. G. Jansen, J. Wang, I. Poser, S. J. van Heeringen, and M. Vermeulen, *Cell Rep.*, **34**, 108705 (2021).
31. X. Chu, T. Sun, Q. Li, Y. Xu, Z. Zhang, L. Lai, and J. Pei, *BMC Bioinformatics*, **23** (2022).
32. R. M. Vernon, P. A. Chong, B. Tsang, T. H. Kim, A. Bah, P. Farber, H. Lin, and J. D. Forman-Kay, *eLife*, **7**, e31486 (2018).
33. H. Cai, R. M. Vernon, and J. D. Forman-Kay, *Biomolecules*, **12**, 1131 (2022).
34. E. A. Paiz, J. H. Allen, J. J. Correia, N. C. Fitzkee, L. E. Hough, and S. T. Whitten, *J. Biol. Chem.*, **297**, 101343 (2021).
35. A. Y. Ibrahim, N. P. Khaodeuanepheng, D. L. Amarasekara, J. J. Correia, K. A. Lewis, N. C. Fitzkee, L. E. Hough, and S. T. Whitten, *J. Biol. Chem.*, **299**, 102801 (2023).