

Plegamiento de Proteínas

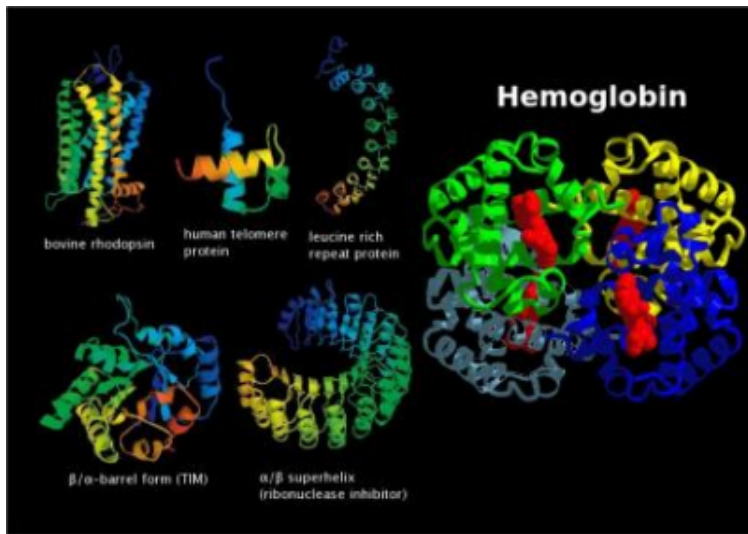
Luis Garreta

Electiva de Bioinformática
MAESTRÍA EN INFORMÁTICA BIOMÉDICA
Universidad del Bosque
Bogotá-Colombia

March 10, 2023

Introducción

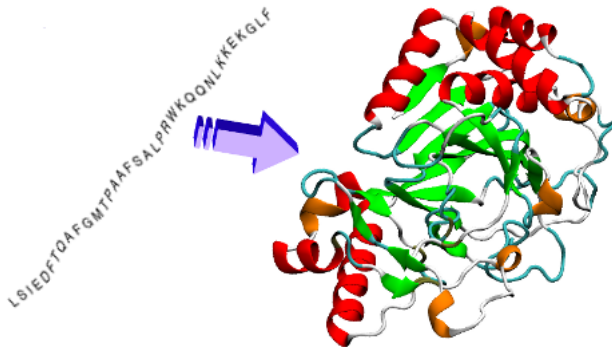
Las proteínas se pliegan en estructuras tridimensionales



https://figshare.com/articles/Hemoglobin_is_a_Flexible_Protein_NMR_Structure_2M6Z_/903674

¿ Cómo se pliegan las proteínas?

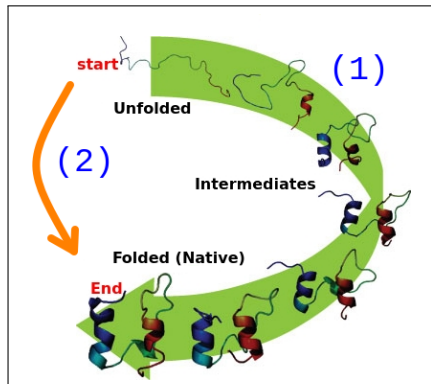
Las proteínas se pliegan desde una secuencia lineal de aminoácidos, sin actividad, hacia una única estructura 3D biológicamente activa.



Proceso de plegamiento de las proteínas

Dos Problemas del plegamiento de proteínas

- ① ¿Cómo se pliegan las proteínas?
 - ▶ Problema del plegamiento de las proteínas
 - ▶ Problema termodinámico
- ① ¿Cuál es la estructura 3D de una proteína?
 - ▶ Problema de la predicción de la estructura de la proteína
 - ▶ Problema computacional.



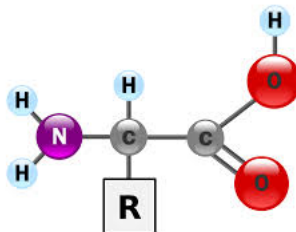
Importancia del plegamiento

- Si entendemos cómo se pliegan las proteínas, podríamos predecir su estructura 3D solo a partir de la información de la secuencia
- La función biológica de una proteína depende de su correcto plegamiento.
- Si una proteína se pliega mal, no será capaz de cumplir con su función biológica.
- Peor aún, si una proteína se pliega mal, puede terminar causando graves enfermedades (Alzheimer, Parkinson, ...):
 - ▶ Puede alcanzar estados aberrantes de agregación, que incluyen la formación de apilamientos amiloides causantes de neuropatías, como ocurre con el llamado prion.

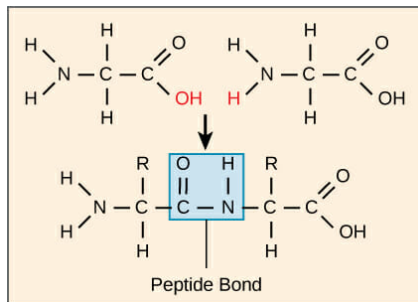
Conceptos

Los aminoácidos se componen de átomos

- Carbono
- Oxígeno
- Hidrógeno
- Nitrógeno

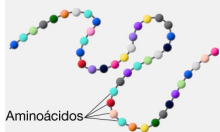


Los aminoácidos están unidos por enlaces peptídicos entre los grupos carboxilo (C-terminal) y amino (N-terminal), que forman la columna vertebral continua de la cadena proteica.

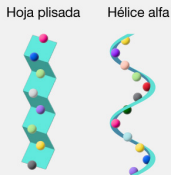


Niveles de organización de las proteínas

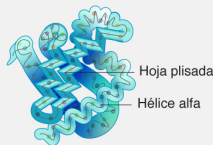
La estructura proteica primaria es la secuencia de una cadena de aminoácidos.



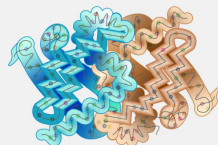
La estructura proteica secundaria se produce cuando la secuencia de aminoácidos se pliega y adopta una forma tridimensional.



La estructura proteica terciaria se produce cuando una proteína madura se pliega sobre sí misma.



La estructura proteica cuaternaria es una proteína que consta de más de una cadena polipeptídica.

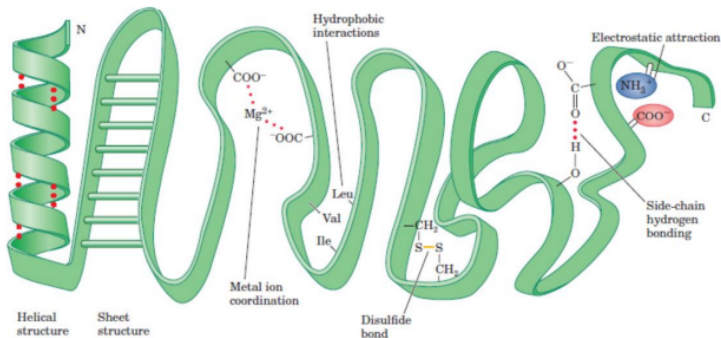
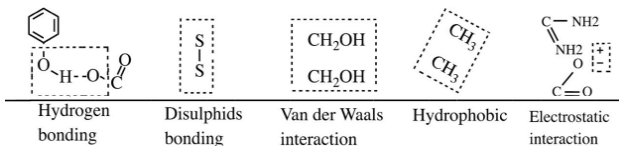


Fuerzas estabilizadoras

Consideraciones sobre el Plegamiento de las Proteínas

- Las proteínas no son estructuras lineales, aunque estén formadas por una cadena lineal de aminoácidos.
- La estructura proteica es clave para su funcionalidad
- Las diferentes propiedades químicas de los aminoácidos son las responsables de las interacciones entre ellos.

Fuerzas que estabilizan el plegamiento



Interacciones covalentes y no-covalentes

- Covalentes:
 - ▶ Puentes de Hidrógeno
 - ▶ Puentes disulfuro
 - ▶ Puentes salinos
- Interacciones no covalentes
 - ▶ Interacciones electrostaticas
 - ▶ Fuerzas de Van der Waals (dipolo — dipolo)
 - ▶ Interacciones hidrofóbicas: fuerza mas importante que dirige el plegamiento de las proteínas.

Las interacciones moleculares que estabilizan una proteína pueden ser alteradas por la temperatura, pH y fuerza iónica

Teorías del Plegamiento de Proteínas

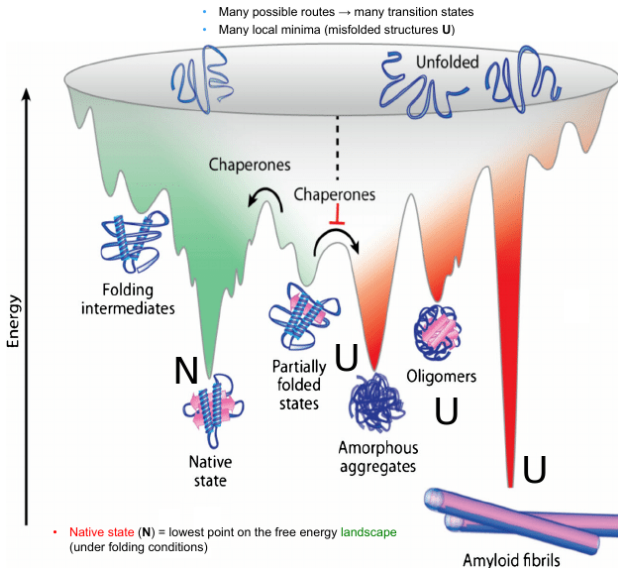
Modelos de Plegamiento de Proteína

- Fundamentos:
 - ▶ Postulado de Anfinsen
 - ▶ Paradoja de Levinthal
- Modelos clásicos:
 - ▶ El modelo de nucleación-propagación.
 - ▶ El modelo de “framework”.
 - ▶ El modelo de difusión-colisión.
 - ▶ El modelo del colapso hidrofóbico.
 - ▶ Concepto del “molten globule” (glóbulo fundido).
- Nueva visión:
 - ▶ Paisaje deplegamiento (*Energy landscape*)

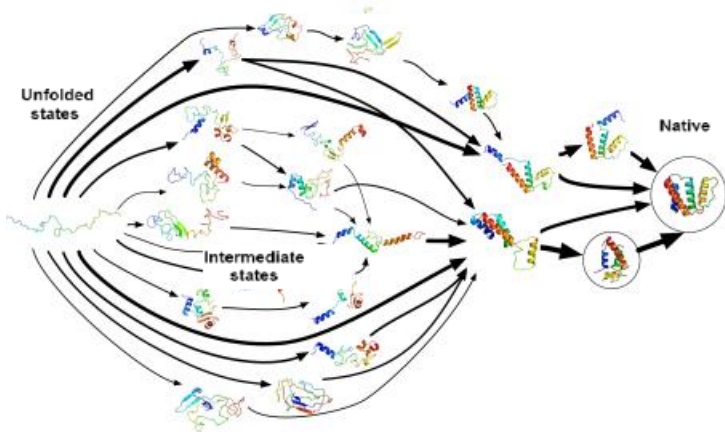
Pasos generales del plegamiento

- ① Formación de estructuras secundarias (α -hélice y B-lámina)
 - ▶ Actúan como núcleos de plegamiento, estabilizando otras regiones ordenadas de la proteína
- ② Formación de dominios
 - ▶ Por agregación cooperativa de distintos núcleos de plegamiento
- ③ Formación del glóbulo fundido o *molten globule*:
 - ▶ En proteínas con varios dominios, dichos dominios se agregan formando un glóbulo fundido
- ④ Transformación del glóbulo fundido en una estructura terciaria
 - ▶ Qué adopta la estructura nativa de una proteína monomérica
 - ▶ Se logra mediante pequeños cambios conformacionales

Paisaje de plegamiento



Rutas de Plegamiento o Protein Folding Pathways



Mal plegamiento

- Una proteína sin estructura nativa:
 - ▶ Es afuncional
 - ▶ Tiende a agregarse con otras cadenas polipeptídicas
 - ▶ Suele ser degradada
 - ▶ Consume recursos celulares (materia y energía)
- Pero si no se desagrega:
 - ▶ Puede realizar una función distinta de la original
 - ▶ Enfermedades

Algunas enfermedades por mal plegamiento

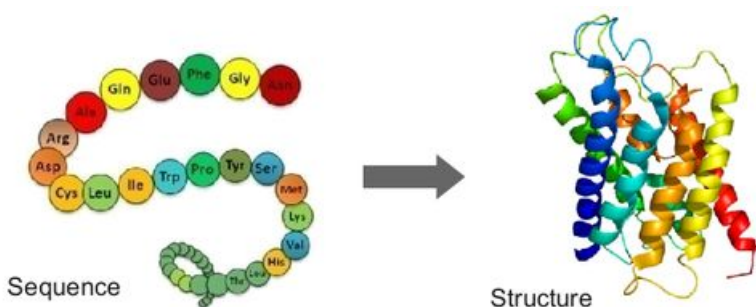
Amiloidosis: conjunto de enfermedades que consiste en el plegamiento anómalo de una proteína precursora que se acaba depositando en forma de estructuras fibrilares en diversos órganos y sistemas (corazón, riñón, hígado, sistema nervioso, etc.) alterando su funcionamiento. Algunas son:

- **Alzheimer:**
Depositos de Beta-amiloide, formando las placas neuríticas
- **Parkinson:** Depositos de Alpha-sinucleína, formando los cuerpos de Lewy
- **Enfisema hereditario:** La alpha1-antitripsina se pliega muy lentamente, por lo que no puede bloquear la acción de su diana, la elastasa, y ésta última destruye el tejido pulmonar
- **Anemia falciforme:** La hemoglobina alterada HbSC promueve la agregación de la Hb dentro de los eritrocitos, disminuyendo su flexibilidad y provocando que adopten una forma de hoz

Computación del plegamiento

Problema de predicción de la estructura proteica

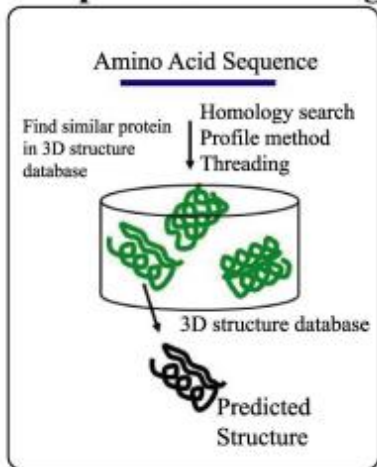
ECómo determinar la estructura nativa de una proteína, dada su secuencia de aminoácidos.



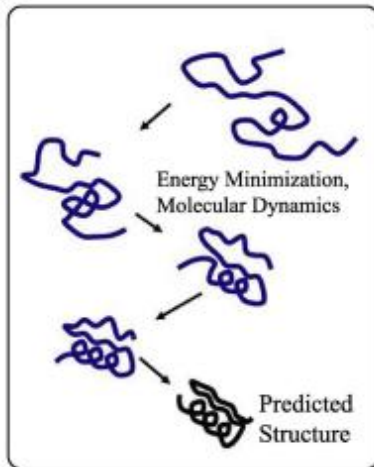
Predicir la estructura de una proteína a partir de su secuencia de aminoácidos sigue siendo un problema sin resolver después de varias décadas de esfuerzos.

Métodos computacionales de predicción de la estructura

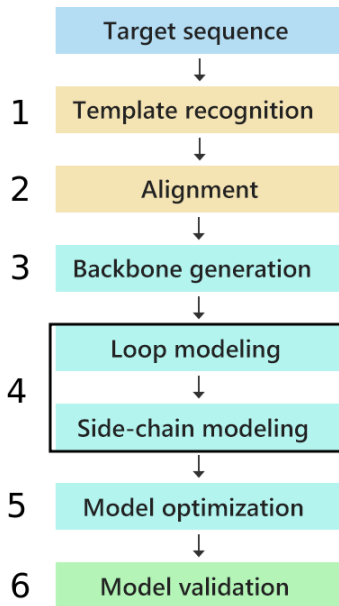
Comparative Modelling



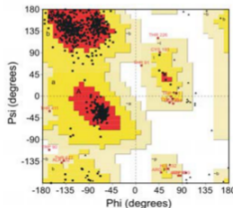
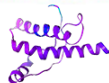
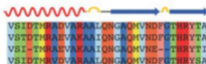
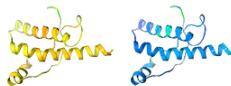
Ab initio Prediction



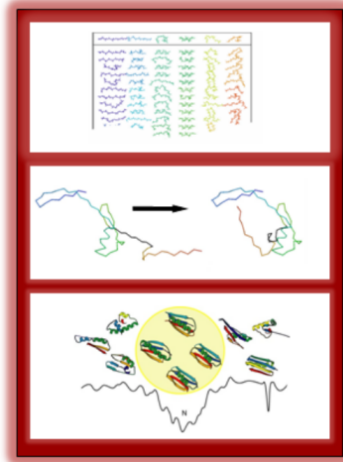
Modelado de homología



VSIDTMRADVARAALQNGAQMVDGFTHRYTA

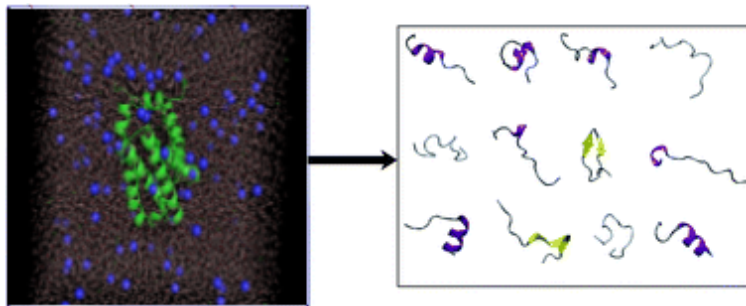


Modelado *de novo* folding



1. Select fragments consistent with local sequence preferences
2. Assemble fragments into models with native-like global properties
3. Identify the best model from the population of decoys

Simulaciones de Dinámica Molecular



Simulating How Proteins Self-Assemble, Or Fold - YouTube:
<https://www.youtube.com/watch?v=gFcp2Xpd29I&t=3s>

Bases de datos y archivos de Proteínas

Protein Data Bank: <https://www.rcsb.org/>

RCSB PDB

Deposit ▾ Search ▾ Visualize ▾ Analyze ▾ Download ▾ Learn ▾ More ▾

MyPDB

RCSB PDB

PROTEIN DATA BANK

140824 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

[Advanced Search](#) | [Browse by Annotations](#)

PDB-101

PDB

EMDataBank

WORLDWIDE PROTEIN DATA BANK

FOUNDED

2000

2018

2020

2022

2024

2026

2028

2030

2032

2034

2036

2038

2040

2042

2044

2046

2048

2050

2052

2054

2056

2058

2060

2062

2064

2066

2068

2070

2072

2074

2076

2078

2080

2082

2084

2086

2088

2090

2092

2094

2096

2098

2100

2102

2104

2106

2108

2110

2112

2114

2116

2118

2120

2122

2124

2126

2128

2130

2132

2134

2136

2138

2140

2142

2144

2146

2148

2150

2152

2154

2156

2158

2160

2162

2164

2166

2168

2170

2172

2174

2176

2178

2180

2182

2184

2186

2188

2190

2192

2194

2196

2198

2200

2202

2204

2206

2208

2210

2212

2214

2216

2218

2220

2222

2224

2226

2228

2230

2232

2234

2236

2238

2240

2242

2244

2246

2248

2250

2252

2254

2256

2258

2260

2262

2264

2266

2268

2270

2272

2274

2276

2278

2280

2282

2284

2286

2288

2290

2292

2294

2296

2298

2300

2302

2304

2306

2308

2310

2312

2314

2316

2318

2320

2322

2324

2326

2328

2330

2332

2334

2336

2338

2340

2342

2344

2346

2348

2350

2352

2354

2356

2358

2360

2362

2364

2366

2368

2370

2372

2374

2376

2378

2380

2382

2384

2386

2388

2390

2392

2394

2396

2398

2400

2402

2404

2406

2408

2410

2412

2414

2416

2418

2420

2422

2424

2426

2428

2430

2432

2434

2436

2438

2440

2442

2444

2446

2448

2450

2452

2454

2456

2458

2460

2462

2464

2466

2468

2470

2472

2474

2476

2478

2480

2482

2484

2486

2488

2490

2492

2494

2496

2498

2500

2502

2504

2506

2508

2510

2512

2514

2516

2518

2520

2522

2524

2526

2528

2530

2532

2534

2536

2538

2540

2542

2544

2546

2548

2550

2552

2554

2556

2558

2560

2562

2564

2566

2568

2570

2572

2574

2576

2578

2580

2582

2584

2586

2588

2590

2592

2594

2596

2598

2600

2602

2604

2606

2608

2610

2612

2614

2616

2618

2620

2622

2624

2626

2628

2630

2632

2634

2636

2638

2640

2642

2644

2646

2648

2650

2652

2654

2656

2658

2660

2662

2664

2666

2668

2670

2672

2674

2676

2678

2680

2682

2684

2686

2688

2690

2692

2694

2696

2698

2700

2702

2704

2706

2708

2710

2712

2714

2716

2718

2720

2722

2724

2726

2728

2730

2732

2734

2736

2738

2740

2742

2744

2746

2748

2750

2752

2754

2756

2758

2760

2762

2764

2766

2768

2770

2772

2774

2776

2778

2780

2782

2784

2786

2788

2790

2792

2794

2796

2798

2800

2802

2804

2806

2808

2810

2812

2814

2816

2818

2820

2822

2824

2826

2828

2830

2832

2834

2836

2838

2840

2842

2844

2846

2848

2850

2852

2854

2856

2858

2860

2862

2864

2866

2868

2870

2872

2874

2876

2878

2880

2882

2884

2886

2888

2890

2892

2894

2896

2898

2900

2902

2904

2906

2908

2910

2912

2914

2916

2918

2920

2922

2924

2926

2928

2930

2932

2934

2936

2938

2940

2942

2944

2946

2948

2950

2952

2954

2956

2958

2960

2962

2964

2966

2968

2970

2972

2974

2976

2978

2980

2982

2984

2986

2988

2990

2992

2994

2996

2998

3000

3002

3004

3006

3008

3010

3012

3014

3016

3018

3020

3022

3024

3026

3028

3030

3032

3034

3036

3038

3040

3042

3044

3046

3048

3050

3052

3054

3056

3058

3060

3062

3064

3066

3068

3070

3072

3074

3076

3078

3080

3082

3084

3086

3088

3090

3092

3094

3096

3098

3100

3102

3104

3106

3108

3110

3112

3114

3116

3118

3120

3122

3124

3126

3128

3130

3132

3134

3136

3138

3140

3142

3144

3146

3148

3150

3152

3154

3156

3158

3160

3162

3164

3166

3168

3170

3172

3174

3176

3178

3180

3182

3184

3186

3188

3190

3192

3194

3196

3198

3200

3202

3204

3206

3208

3210

3212

3214

3216

3218

3220

3222

3224

3226

3228

3230

3232

3234

3236

3238

3240

3242

3244

3246

3248

3250

3252

3254

3256

3258

3260

3262

3264

3266

3268

3270

3272

3274

3276

3278

3280

3282

3284

3286

3288

3290

3292

3294

3296

3298

3300

3302

3304

3306

3308

3310

3312

3314

3316

3318

3320

3322

3324

3326

3328

3330

3332

3334

3336

3338

3340

3342

3344

3346

3348

3350

3352

3354

3356

3358

3360

3362

3364

3366

3368

3370

3372

3374

3376

3378

3380

3382

3384

3386

3388

3390

3392

3394

3396

3398

3400

3402

3404

3406

3408

3410

3412

3414

3416

3418

3420

3422

3424

3426

3428

3430

3432

3434

3436

3438

3440

3442

3444

3446

3448

3450

3452

3454

3456

3458

3460

3462

3464

3466

3468

3470

3472

3474

3476

3478

3480

3482

3484

3486

3488

3490

3492

3494

3496

3498

3500

3502

3504

3506

3508

3510

3512

3514

3516

3518

3520

3522

3524

3526

3528

3530

3532

3534

3536

3538

3540

3542

3544

3546

3548

3550

3552

3554

3556

3558

3560

3562

3564

3566

3568

3570

3572

3574

3576

3578

3580

3582

3584

3586

3588

3590

3592

3594

3596

3598

3600

3602

3604

3606

3608

3610

3612

3614

3616

3618

3620

3622

3624

3626

3628

3630

3632

3634

3636

3638

3640

3642

3644

3646

3648

3650

3652

3654

3656

3658

3660

3662

3664

3666

3668

3670

3672

3674

3676

3678

3680

3682

3684

3686

3688

3690

3692

3694

3696

3698

3700

3702

3704

3706

3708

3710

3712

3714

3716

3718

3720

3722

3724

3726

3728

3730

3732

3734

3736

3738

3740

3742

3744

3746

3748

3750

3752

3754

3756

3758

3760

3762

3764

3766

3768

3770

3772

3774

3776

3778

3780

3782

3784

3786

3788

3790

3792

3794

3796

3798

3800

3802

3804

3806

3808

3810

3812

3814

3816

3818

3820

3822

3824

3826

3828

3830

3832

3834

3836

3838

3840

3842

3844

3846

3848

3850

3852

3854

3856

3858

3860

3862

3864

3866

3868

3870

3872

3874

3876

3878

3880

3882

3884

3886

3888

3890

3892

3894

3896

3898

3900

3902

3904

3906

3908

3910

3912

3914

3916

3918

3920

3922

3924

3926

3928

3930

3932

3934

3936

3938

3940

3942

3944

3946

3948

3950

3952

3954

3956

3958

3960

3962

3964

3966

3968

3970

3972

3974

3976

3978

3980

3982

3984

3986

3988

3990

3992

3994

3996

3998

4000

4002

4004

4006

4008

4010

4012

4014

4016

4018

4020

4022

4024

4026

4028

4030

4032

4034

4036

4038

4040

4042

4044

4046

4048

4050

4052

4054

4056

4058

4060

4062

4064

4066

4068

4070

4072

4074

4076

4078

4080

4082

4084

4086

4088

4090

4092

4094

4096

4098

4100

4102

4104

4106

4108

4110

4112

4114

4116

4118

4120

4122

4124

4126

4128

4130

4132

4134

4136

4138

4140

4142

4144

4146

4148

4150

4152

4154

4156

4158

4160

4162

4164

4166

4168

4170

4172

4174

4176

4178

4180

4182

4184

4186

4188

4190

4192

4194

4196

4198

4200

4202

4204

4206

4208

4210

4212

4214

4216

4218

4220

4222

4224

4226

4228

4230

4232

4234

4236

4238

4240

4242

4244

4246

4248

4250

4252

4254

4256

4258

4260

4262

4264

4266

4268

4270

4272

4274

4276

4278

4280

4282

4284

4286

4288

4290

4292

4294

4296

4298

4300

4302

4304

4306

4308

4310

4312

4314

4316

4318

4320

4322

4324

4326

4328

4330

4332

4334

4336

4338

4340

4342

4344

4346

4348

4350

4352

4354

4356

4358

4360

4362

4364

4366

4368

4370

4372

4374

4376

4378

4380

4382

4384

4386

4388

4390

4392

4394

4396

4398

4400

4402

4404

4406

4408

4410

4412

4414

4416

4418

4420

4422

4424

4426

4428

4430

4432

4434

4436

4438

4440

4442

4444

4446

4448

4450

4452

4454

4456

4458

4460

4462

4464

4466

4468

4470

4472

4474

4476

4478

4480

4482

4484

4486

4488

4490

4492

4494

4496

4498

4500

4502

4504

4506

4508

4510

4512

4514

4516

4518

4520

4522

4524

4526

4528

4530

4532

4534

4536

4538

4540

4542

4544

4546

4548

4550

4552

4554

4556

4558

4560

4562

4564

4566

4568

4570

4572

4574

4576

4578

4580

4582

4584

4586

4588

4590

4592

4594

4596

4598

4600

4602

4604

4606

4608

4610

4612

4614

4616

4618

4620

4622

4624

4626

4628

4630

4632

4634

4636

4638

4640

4642

4644

4646

4648

4650

4652

4654

4656

4658

4660

4662

4664

4666

4668

4670

4672

4674

4676

4678

4680

4682

4684

4686

4688

4690

4692

4694

4696

4698

4700

470

Protein Data Bank

- El Protein Data Bank (PDB) es una base de datos donde se almacenan las estructuras cuya estructura tridimensional (es decir, sus coordenadas atómicas) ha sido resuelta.
- Estos datos, generalmente obtenidos por Cristalografía de rayos X o Resonancia Magnética Nuclear, son enviados por biólogos y bioquímicos de todo el mundo. Están bajo el dominio público y pueden ser usados libremente.

Archivo de secuencias de aminoácidos: Formato fasta

Formato Fasta:

- El encabezado en la primera línea
- Seguido de la secuencia

```
>sp|P26239|BCHI_RHOCB Magnesium-chelatase 38 kDa subunit OS=Rhodobacter capsulatus
MTTAVARLQPSASGAKTRPVFPFSAIVGQEDMKLALLLTAVDPGIGGVLVFGDRGTGKST
AVRALAALLPEIEAVEGCPVSSPNVEMIPDWATVLSTNVIRKPTPVVDLPLGVSEDRVVG
ALDIERAISKGEKAFEPGLLARANGYLYIDECNLLEDHIVDLLLDVAQSGENVVERDGL
SIRHPARFVLVSGCNPEEGDLRPQLLDRFGLSVEVLSPRDVETRVEVIRRRDTYDADPKA
FLEEWRPKMDIRNQILEARERLPKVEAPNTALYDCAALCIALGSDGLRGELTLLRSARA
LAALEGATAVGRDHLKRVATMALSHRLRRDPLDEAGSTARVARTVEETLP
```

Archivos de Estructuras de Proteínas: Formato PDB

Atomic Coordinates: PDB Format

		Amino Acid		Chain name		Sequence Number		-----Coordinates-----		
		Element						X	Y	Z
ATOM	1	N	ASP	L	1			4.060	7.307	5.186
ATOM	2	CA	ASP	L	1			4.042	7.776	6.553
ATOM	3	C	ASP	L	1			2.668	8.426	6.644
ATOM	4	O	ASP	L	1			1.987	8.438	5.606
ATOM	5	CB	ASP	L	1			5.090	8.827	6.797
ATOM	6	CG	ASP	L	1			6.338	8.761	5.929
ATOM	7	OD1	ASP	L	1			6.576	9.758	5.241
ATOM	8	OD2	ASP	L	1			7.065	7.759	5.948

\\
Element position within amino acid

