Work Report

贾宁欣

09-11-2021

Post-processing procedure

Apoc

Problem 1: Apoc 默认只识别残基数 ≥ 10的口袋

solution-1 : 将默认参数设置为0

$$\text{PS-score} = (S + s_0)/(1 + s_0) \quad (1)$$

$$S = \frac{1}{L_Q} \max_{sup} \left[\sum_{i=1}^{N_a} p_i r_i / (1 + d_i^2/d_0^2) \right] \quad (2)$$

$$p_i = \begin{cases} 1 & if\ \theta_i \le \pi/3 \\ \max(0.1, 0.5 + \cos\theta_i) & if\ \theta_i > \pi/3 \end{cases} \quad (3)$$

$$r_i = \max(0.8, \delta(a_i^Q, a_i^T i)) \quad (4)$$

$$s_0 \equiv 0.23 - 12/L_Q^{1.88}$$

L_Q: length of query pocket

当$L_Q < 7.876$，$s_0 < 0$，PS-score < 0

solution-2 : 仍将默认参数设置为0，但调换口袋输入 Apoc 的顺序，使PLDB库中的模板口袋的残基数作为L_Q，使PS-score为负的口袋舍去。

```
1. To compare two structures,
   > ../bin/apoc 1ha3A.pdb 3ec1A.pdb

2. To compare a list of templates against a query
   > ../bin/apoc -lt templ.lst 1yr8A.pdb
```

solution-3 : PPS-align

Post-processing procedure

Apoc (solution-2)

3j8yK-6h2jB.txt

```
----------NFIVLDKYIKAEPTGDSYQSESDLERELIQDLRNQGYEFISVKSQSAMLANVREQLQNLNGVVFNDSEWRRFTEQY

Scoring parameters: normalization length = 990,  d0 =10.496
Alignment search mode: sequential
Best alignment search: initial = 31
Sequence similarity (BLOSUM62): sum = -125, sum_pos = 128
Running time: 6.69350 seconds


>>>>>>>>>>>>>>>>>>>>>>>>   Pocket alignment   <<<<<<<<<<<<<<<<<<<<<<<<<
Structure 1: 3j8yK.pdb                      Length =   8 AAs, Pocket:3j8yK_ATP_BS01
Structure 2: 6h2jB.pdb                      Length =  17 AAs, Pocket:6h2jB_ATP_BS01

PS-score = 0.41113, P-value = 0.7430E-005, Z-score = 11.810
Number of aligned residues  = 7
RMSD =  0.63, Seq identity  = 0.714

 -------- rotation matrix to rotate Chain-1 to Chain-2 ------
 i          t(i)         u(i,1)         u(i,2)         u(i,3)
 1   -143.6381043061  -0.1285802165   0.9889620664   0.0736285219
 2     43.4071317242   0.8376378529   0.0685606768   0.5419061366
 3      8.5704844086   0.5308765914   0.1313524454  -0.8372076086

 ****** Match List ******
 Index Ch1 Resid1  AA1 Ch2 Resid2  AA2 Distance Cos(theta)
    1    K   83     T   A  282     T   0.180      0.945   *
    2    K   84     S   A  283     G   0.224      0.272
    3    K   85     S   A  284     S   0.250      0.999   *
    4    K   86     G   A  285     G   0.165      1.000   *
    5    K   87     K   A  286     K   0.632      0.952   *
    6    K   88     T   A  287     T   0.533      0.896   *
    7    K   89     H   A  288     L   1.843      0.991

Scoring parameters: normalization length = 17,  d0 = 1.103
Alignment search mode: non-sequential
Best alignment search: initial = 12
Sequence similarity (BLOSUM62): sum = 22, sum_pos = 25
Running time: 0.02180 seconds
```

	A	B	C	D	E	F	G	H	I	J
1	prd-ABRs	13	83	84	85	86	87	88	89	
600	5xd2A_ATP_BS02		0.25148	0.25148	0.25148		0.25148	0.25148		
601	5xmiB_ATP_BS01	0.49394		0.49394	0.49394	0.49394	0.49394	0.49394	0.49394	
602	5xvuC_ATP_BS01					0.29596	0.29596	0.29596	0.29596	
603	5y0nB_ATP_BS01	0.26687		0.26687	0.26687	0.26687	0.26687	0.26687	0.26687	
604	5yecA_ATP_BS01		0.28292	0.28292	0.28292	0.28292	0.28292	0.28292		
605	5yecC_ATP_BS01	0.27723			0.27723	0.27723	0.27723	0.27723	0.27723	
606	5yh3B_ATP_BS01			0.22662		0.22662	0.22662	0.22662	0.22662	
607	5yh3B_ATP_BS02	0.14451				0.14451		0.14451	0.14451	
608	5yudA_ATP_BS01	0.47663	0.47663	0.47663		0.47663	0.47663	0.47663	0.47663	
609	5z68C_ATP_BS01	0.47692	0.47692	0.47692	0.47692	0.47692	0.47692	0.47692	0.47692	
610	5zxdA_ATP_BS01		0.61256	0.61256		0.61256	0.61256	0.61256	0.61256	
611	5zxdA_ATP_BS02	0.48799	0.48799	0.48799		0.48799	0.48799	0.48799	0.48799	
612	5zxdB_ATP_BS01	0.50214	0.50214	0.50214		0.50214	0.50214	0.50214	0.50214	
613	5zxdB_ATP_BS02			0.53003		0.53003	0.53003	0.53003	0.53003	
614	6bcqA_ATP_BS01	0.20871		0.20871		0.20871	0.20871		0.20871	
615	6bhuA_ATP_BS01	0.36644	0.36644	0.36644	0.36644	0.36644	0.36644	0.36644	0.36644	
616	6bhuA_ATP_BS02	0.4016	0.4016	0.4016	0.4016	0.4016	0.4016	0.4016	0.4016	
617	6blkA_ATP_BS01	0.34635		0.34635	0.34635	0.34635	0.34635	0.34635	0.34635	
618	6blkC_ATP_BS01	0.28438	0.28438	0.28438	0.28438	0.28438		0.28438	0.28438	
619	6c0vA_ATP_BS01	0.41029	0.41029	0.41029		0.41029	0.41029	0.41029	0.41029	
620	6c0vA_ATP_BS02		0.41619	0.41619		0.41619	0.41619	0.41619	0.41619	
621	6c3oG_ATP_BS02		0.45473	0.45473	0.45473	0.45473	0.45473	0.45473	0.45473	
622	6c3pA_ATP_BS01	0.33954	0.33954	0.33954		0.33954	0.33954			
623	6c3pH_ATP_BS02		0.47248	0.47248	0.47248	0.47248	0.47248	0.47248	0.47248	
624	6cn2A_ATP_BS01		0.20459		0.20459	0.20459	0.20459			
625	6cp6A_ATP_BS01	0.48678	0.48678	0.48678	0.48678	0.48678	0.48678	0.48678	0.48678	
626	6ctaA_ATP_BS01		0.27183	0.27183		0.27183	0.27183	0.27183	0.27183	
627	6fhsJ_ATP_BS01	0.33588	0.33588		0.33588	0.33588	0.33588	0.33588	0.33588	
628	6fkfA_ATP_BS01	0.546	0.546	0.546	0.546	0.546	0.546	0.546	0.546	
629	6fmlJ_ATP_BS01	0.32373	0.32373	0.32373	0.32373	0.32373	0.32373	0.32373	0.32373	
630	6h2jB_ATP_BS01		0.41113	0.41113	0.41113	0.41113	0.41113	0.41113	0.41113	
631										
632	SUM	122.50263	147.1991	169.88294	159.38513	198.29172	180.75158	179.59431	175.12072	
633	SUN/629	0.19475776	0.23402083	0.27008417	0.25339448	0.31524916	0.28736340	0.28552355	0.27841132	
634										
635										

3j8yK

3j8yK.txt

```
13  0.19475775834658182
83  0.23402082670906207
84  0.2700841653418125
85  0.25339448330683645
86  0.315249157392687
87  0.28736340222575524
88  0.2855235453100159
89  0.278411319554849
```

Post-processing procedure

Apoc (solution-2)

x.txt

423	0	1	0.10893435200000004
424	1	1	0.10761040800000012
425	1	1	0.10752268800000006
426	1	1	0.10633952861952856
427	0	1	0.10613181355932201
428	0	1	0.10602848101265822
429	1	1	0.10574689774696704
430	0	1	0.10508508661417326
431	1	1	0.10380536000000008
432	1	1	0.10312101761252447
433	1	1	0.10273436026936025
434	0	1	0.10265109589041099
435	0	1	0.10219194127243066
436	1	1	0.10098785365853655
437	0	1	0.10049083464566932
438	1	1	0.10035724505327238
439	1	1	0.10034206400000002
440	1	1	0.0963954400000001
441	0	1	0.09350557911908641
442	0	1	0.09182639921722112
443	1	1	0.0908738983050847
444	0	1	0.09065104404567698
445	1	1	0.09051852000000005
446	1	1	0.08670650455927054
447	1	1	0.08528531810766724
448	0	1	0.08492227272727274
449	1	1	0.08314390572390573
450	0	1	0.08292618045112789
451	0	1	0.08141072440944883
452	0	1	0.08086618444846294
453	1	1	0.07426296360485266
454	0	1	0.0723707922535212
455	0	1	0.07185556836902798
456	0	1	0.06424515370705246
457	1	1	0.05278832
458	1	1	0.05221112115732368
459	1	1	0.03542597600000001
460	0	1	0.027276248153618902
461	1	1	0.024658504
462	0	1	0.00342386301369863
463	0	1	0.0032291400832177536
464	0	1	0.0013099167822468792
465			

Post-processing		TP	FP	TN	FN	Sen	Spe	Pre	ACC	MCC
without		386	78	14081	288	0.5727	0.9945	0.8319	0.9753	**0.67858**
with	0.10	375	64	14095	299	0.5564	0.9955	0.8542	0.9755	**0.67819**
	0.09	378	67	14092	296	0.5608	0.9953	0.8494	0.9755	**0.67892**
	0.08	381	71	14088	293	0.5652	0.9950	0.8429	0.9755	**0.67886**
	0.07	382	73	14086	292	0.5668	0.9948	0.8396	0.9754	**0.67831**

1\ 数据量太少。 388

2\ 只是从寻找FP的角度，还没有考虑FN（没能预测出的绑定残基）。

Post-processing procedure

PSS-align (solution-3)