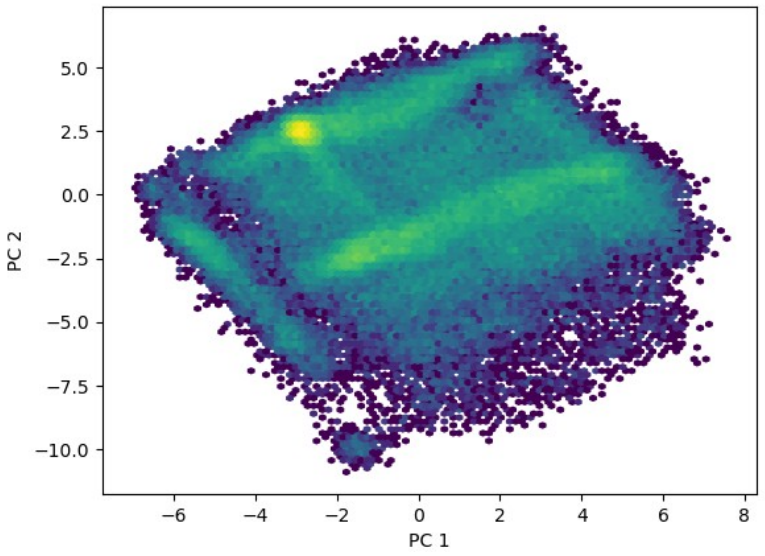
Analysis of 100 us COVID19 Protease Simulation (D. Shaw)

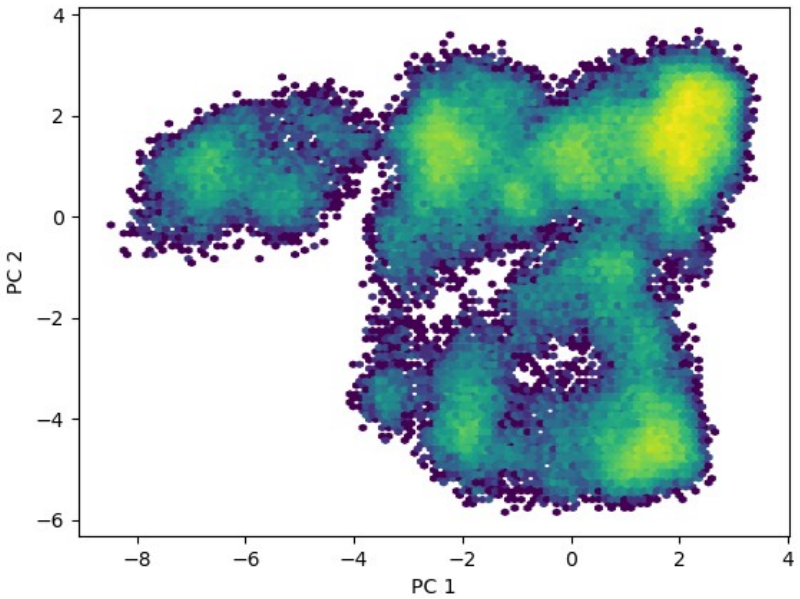
PCA (Cartesian coordinates)



Cumulative variance

```
[0.14949101 0.28873941 0.37822441 0.44074239 0.49958061 0.54105917
 0.57501158 0.60025959 0.62121608 0.64103417 0.66004669 0.67695827
 0.69276077 0.70680536 0.71976451 0.73078621 0.74171396 0.7507201
 0.75864976 0.7658033  0.77228364 0.7783319  0.78377882 0.78906058
 0.79398807 0.79850847 0.80292197 0.8069265  0.81089965 0.81467107
 0.81837613 0.82191574 0.82532047 0.82850123 0.83156683 0.83449962
 0.83734823 0.84008326 0.84275464 0.8453097  0.84773026 0.85011669
 0.85246877 0.85468534 0.85678273 0.85883754 0.86085996 0.86280329
 0.86465511 0.86648884 0.86823753 0.86993206 0.87160019 0.87320406
 0.87476686 0.87630222 0.87781812 0.87931937 0.88075833 0.88214786
 0.88351772 0.88484692 0.88614877 0.88743752 0.88870022 0.88993407
 0.8911352  0.89228196 0.89341132 0.89451349 0.89560031 0.8966281
 0.89764093 0.89863389 0.89961663 0.90058826 0.90154021 0.90247616
 0.90339632 0.90430016 0.90519611 0.90607174 0.90693521 0.90778644
 0.90862799 0.90944651 0.91024668 0.91103394 0.91181703 0.91257827
 0.91332644 0.91404989 0.91476937 0.915479   0.91616554 0.91684653
...
```

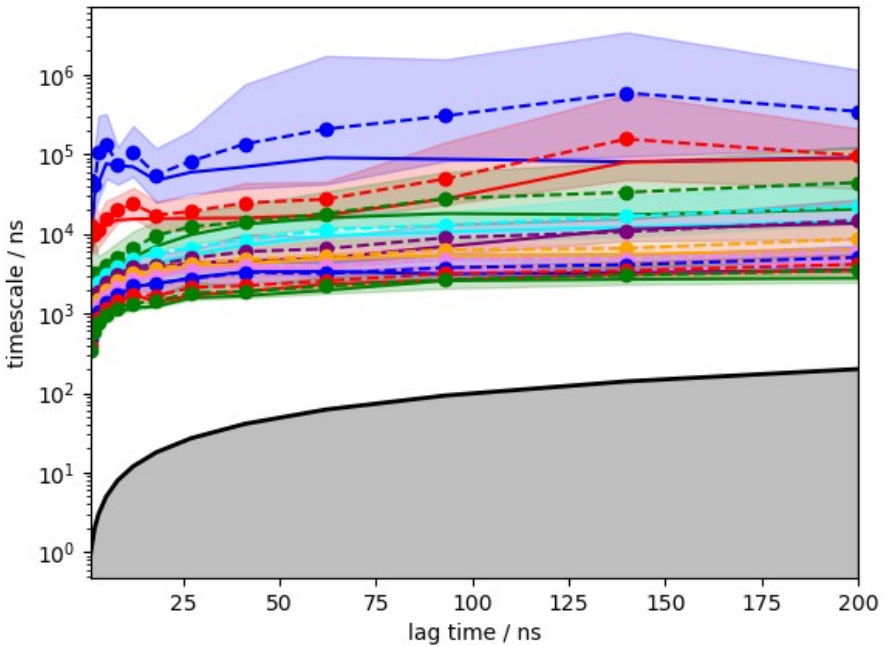
PCA (torsions)



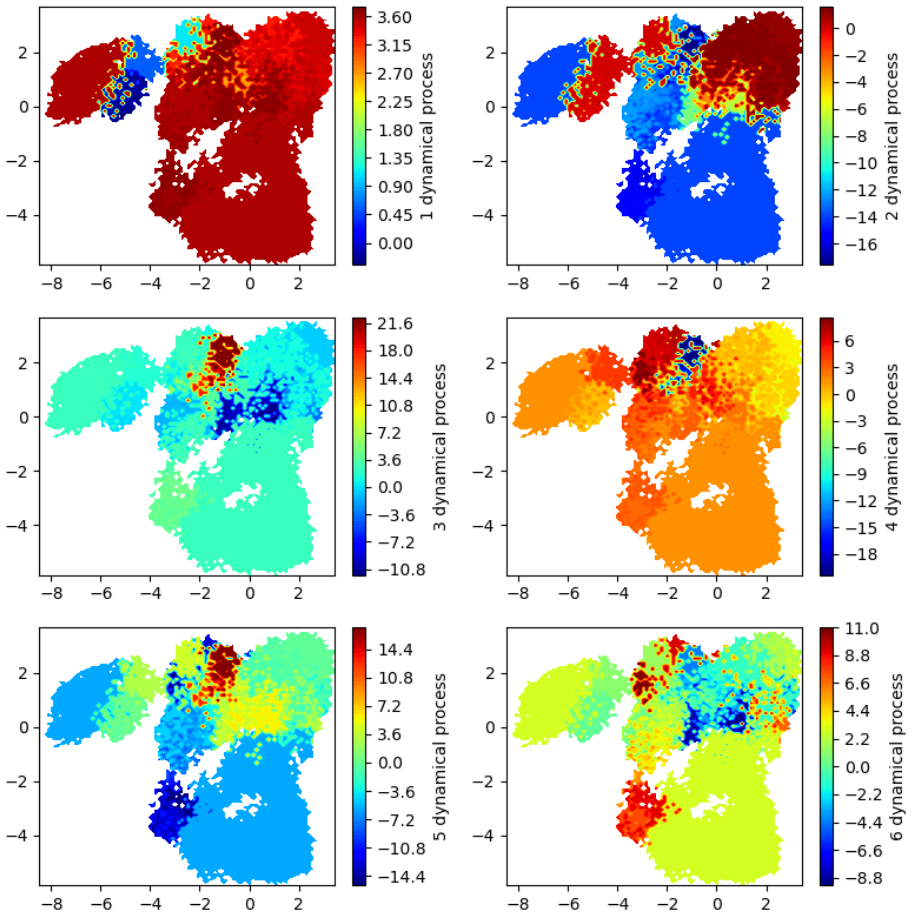
Cumulative variance

```
[0.04965511 0.09773939 0.12330422 0.14485369 0.16125899 0.17468184
 0.18697304 0.19797596 0.20833532 0.21733629 0.22592203 0.23362818
 0.24128645 0.24826598 0.25513665 0.26170803 0.26795005 0.2736372
 0.27915599 0.28451322 0.28975052 0.29489282 0.29990529 0.30461041
 0.30929811 0.3138274  0.31828584 0.32264621 0.32684629 0.33093289
 0.3348784  0.33876657 0.34253285 0.34621222 0.34975542 0.35328082
 0.3566916  0.36001522 0.36327166 0.36647489 0.36964233 0.37277217
 0.3758743  0.37894509 0.38199123 0.38501577 0.38802494 0.39098619
 0.39389461 0.39674652 0.39957576 0.40236994 0.40514043 0.40781289
 0.41045756 0.4130646  0.41562871 0.41818674 0.42069878 0.42318101
 0.42564137 0.42806463 0.43047484 0.43285193 0.43519803 0.43753797
 0.43986443 0.4421554  0.44443891 0.44668629 0.44892067 0.45112607
 0.45332493 0.45548438 0.45760683 0.45971737 0.46179674 0.46387003
 0.4659237  0.46795864 0.46996481 0.47195886 0.47392851 0.47587637
 0.47780923 0.47971399 0.48160608 0.48348704 0.48535184 0.48721088
 0.48905261 0.49086471 0.49266919 0.49446166 0.49625198 0.49803827
 0.49979765 0.50155143 0.50329258 0.5050243  0.50674937 0.50845097
 0.51015115 0.51183365 0.51351241 0.51516932 0.51682389 0.51846917
 0.5200898  0.52170081 0.52329729 0.52488708 0.52646871 0.52804568
 0.52960657 0.53115492 0.53268785 0.53421964 0.53574767 0.53725588
 0.53876116 0.54025324 0.54173861 0.5432124  0.54467903 0.54612953
 0.54757502 0.54901339 0.55044677 0.55186658 0.55327763 0.55467698
 0.55607149 0.55744983 0.55881791 0.56018284 0.56153805 0.56289147
 0.56422911 0.56556114 0.56688938 0.56821074 0.56953    0.57084343
 0.57215305 0.57345291 0.57474798 0.57604167 0.57732985 0.57861545
 0.57989005 0.5811619  0.58242894 0.58368987 0.584945   0.58619657
 0.58744238 0.58868433 0.58992252 0.59115577 0.59238397 0.59360663
...
```

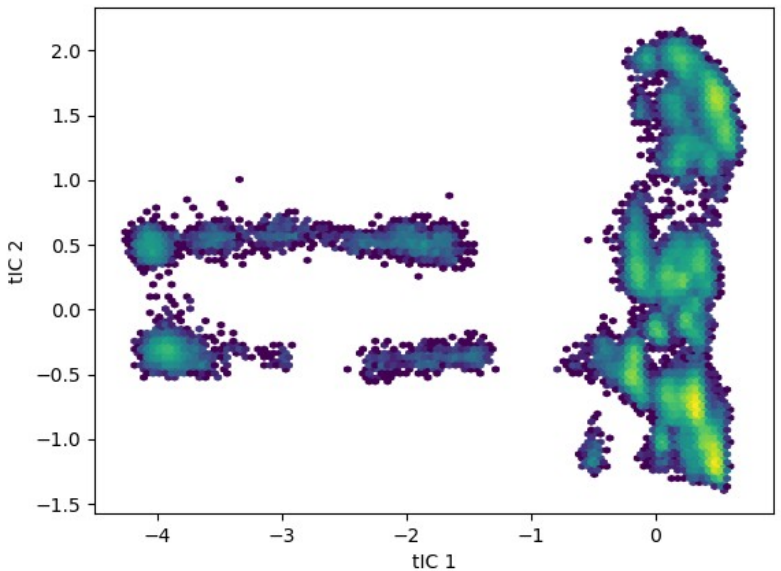
MSM



Pretty converged timescales, but clustering is not really kinetic



TICA (torsions)

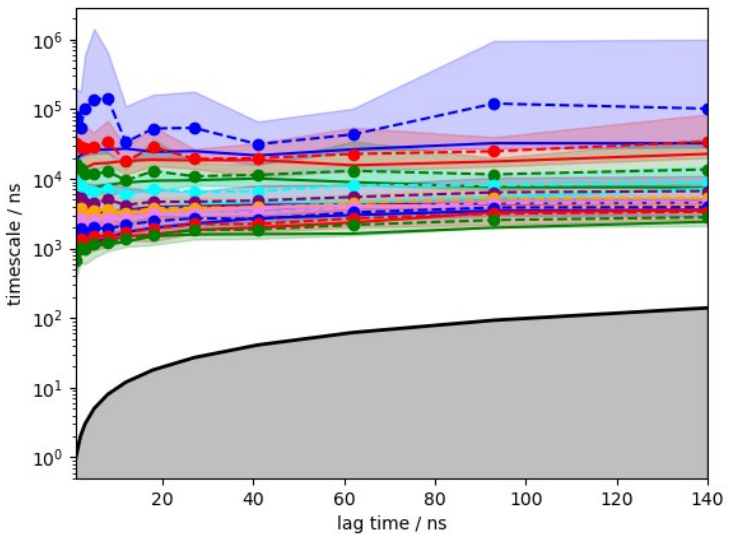


Kinetic cumulative variance

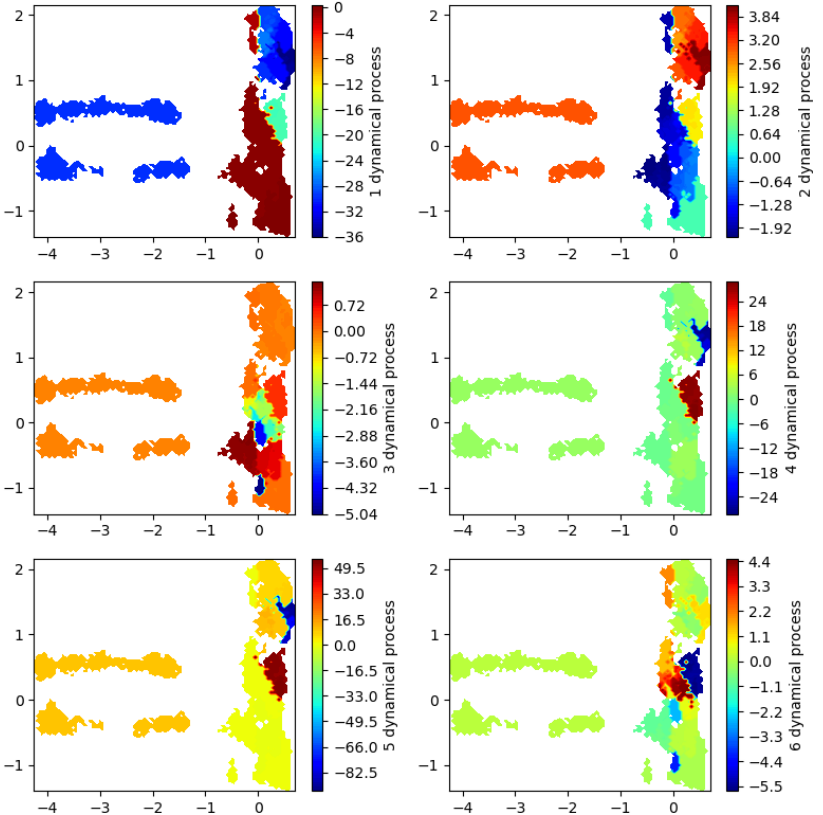
```
[0.00660235 0.01318902 0.01974568 0.02628765 0.03278207 0.03921978
 0.04561639 0.05199281 0.05831633 0.06462561 0.07086871 0.07708092
 0.08328166 0.08946244 0.09562266 0.10175021 0.10779634 0.11379698
 0.11977649 0.12575405 0.13171508 0.13763417 0.14353959 0.14940734
 0.1552302  0.16102881 0.16679169 0.1725188  0.17822727 0.18388802
 0.18950382 0.19505632 0.20057929 0.20609065 0.21155146 0.21700705
 0.22242796 0.22782874 0.23319317 0.23852487 0.2438337  0.24909524
 0.25435222 0.25958392 0.26476031 0.2698996  0.27501665 0.28011999
 0.28518    0.29018896 0.29518021 0.30015102 0.3050746  0.30996516
 0.31480169 0.3196266  0.32440261 0.32915331 0.33389704 0.33861396
 0.34328691 0.34792993 0.35249289 0.35704201 0.36154778 0.36602938
 0.37047855 0.37491355 0.3793214  0.38366966 0.38795266 0.39220026
 0.39640695 0.40059541 0.40476717 0.40889958 0.41300184 0.41707147
 0.42110827 0.4251265  0.42910114 0.43305908 0.43697366 0.44087793
 0.44474634 0.448589   0.45238898 0.45614431 0.45985984 0.46356342
 0.46724054 0.47087443 0.47449305 0.47805554 0.48156495 0.48505465
 0.48853295 0.49196239 0.49534677 0.49869363 0.50201741 0.50528864
 0.50853727 0.51175296 0.51493502 0.51809213 0.52120659 0.52430267
 0.52737909 0.5304285  0.53340076 0.53633065 0.53921265 0.54207159
 0.54490197 0.54772115 0.55051358 0.55327988 0.55603415 0.55874511
 0.56143856 0.56406778 0.56668424 0.56927998 0.57185238 0.57441299
 0.57692878 0.57941394 0.58186234 0.58430072 0.58672468 0.58911697
 0.5914927  0.59385686 0.59615982 0.59844571 0.60071914 0.60297036
 0.6051944  0.60740474 0.60960266 0.61179526 0.61393448 0.61606077
 0.6181612  0.62024793 0.62230704 0.62436001 0.62638365 0.62839022
 0.63037444 0.63234457 0.63430173 0.63623489 0.6381577  0.64006008
 0.64194356 0.64381776 0.645668   0.64750576 0.64931802 0.65111751
 0.65289203 0.65464979 0.65639929 0.65813    0.65985208 0.66156005
 0.66325411 0.66492447 0.66659413 0.66824247 0.66987726 0.67150349
...
```

MSM 2

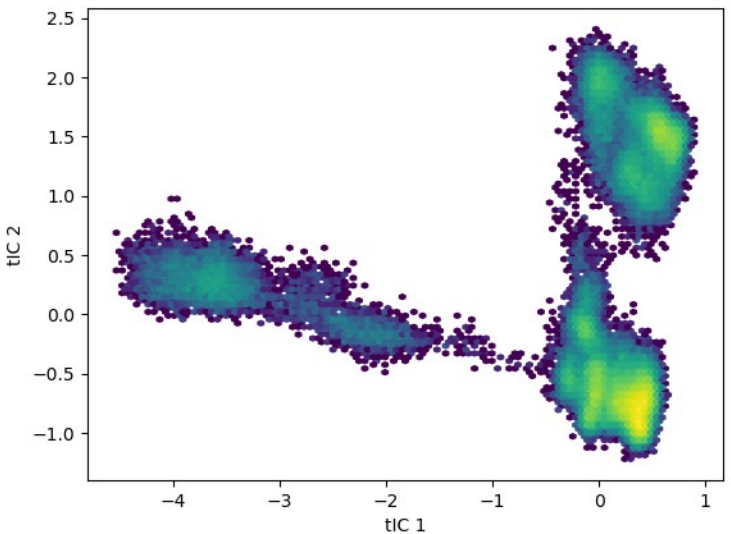
Cluster in 5 tICA dimensions instead of all dimensions giving 95% cumulative kinetic variance



Good convergence! Slower processes slightly undersampled



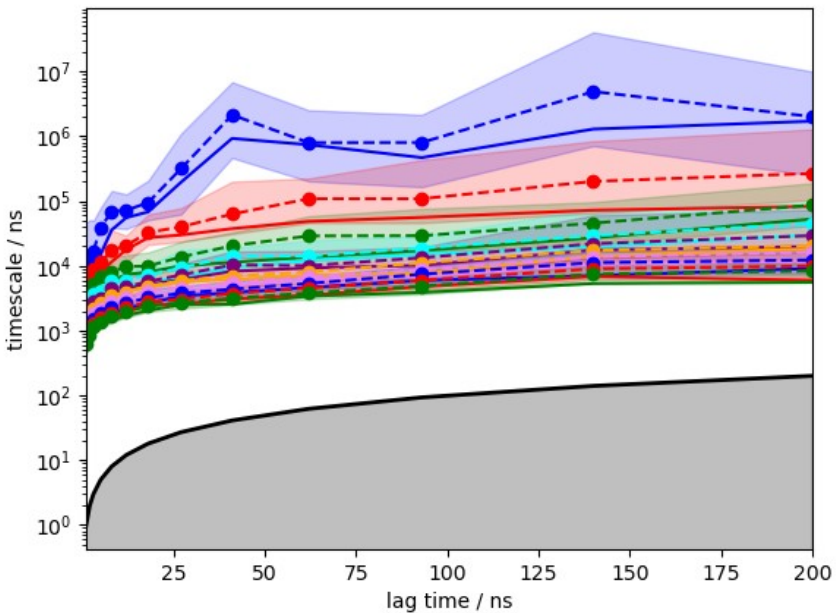
TICA (Cartesian coordinates)



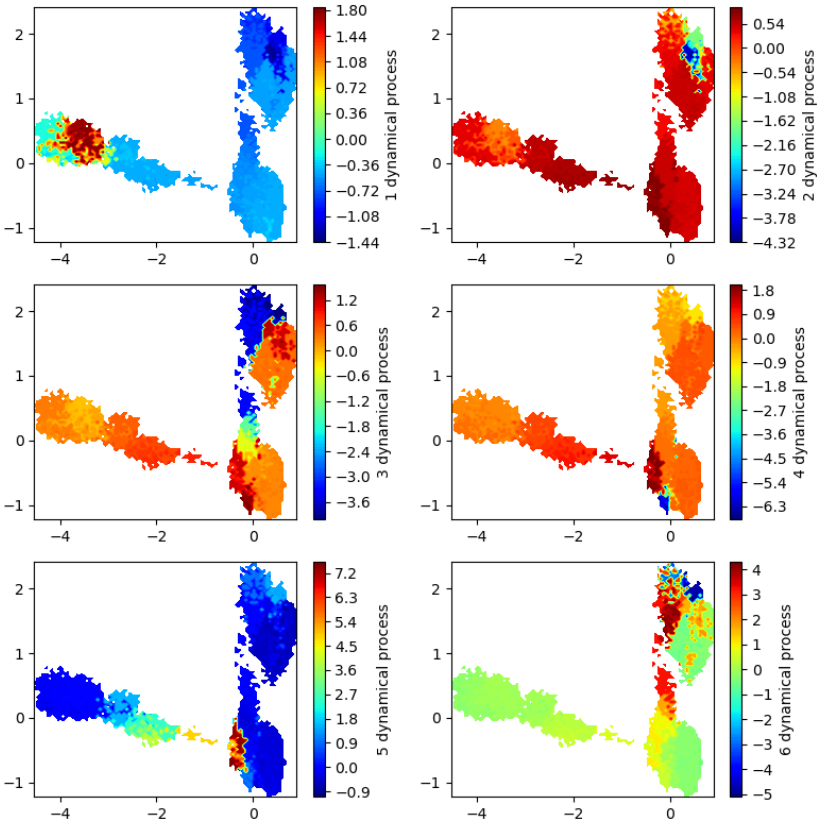
Kinetic cumulative variance

```
[0.00774162 0.01545096 0.02309604 0.03064542 0.03814186 0.04559529
 0.05296469 0.06026234 0.06749348 0.07469994 0.08188433 0.0889776
 0.09600162 0.10299255 0.10992592 0.11680483 0.12360073 0.13038708
 0.13713693 0.14379869 0.15042578 0.15699551 0.16353611 0.17000764
 0.17639302 0.18273612 0.18903791 0.19531739 0.20152919 0.20766133
 0.21376353 0.21985752 0.22588505 0.23185568 0.23778964 0.24367535
 0.24949654 0.25527653 0.26096581 0.26663931 0.2723055  0.27786508
 0.28337131 0.28885988 0.29429868 0.29966563 0.30497442 0.31024609
 0.31548587 0.32067284 0.3257942  0.33088814 0.33589825 0.34088409
 0.34584914 0.35076191 0.35562932 0.36042873 0.36521478 0.36995457
 0.37463934 0.37927329 0.38386044 0.3884408  0.39295856 0.3974283
 0.40185692 0.40623548 0.41059232 0.41491894 0.41917283 0.42339835
 0.4276028  0.43175796 0.43590091 0.44001233 0.44410795 0.44814485
 0.45214284 0.45610102 0.46003191 0.46391773 0.46776617 0.47154885
 0.47531761 0.47903303 0.48272652 0.48639608 0.49003667 0.49363135
 0.49717509 0.50069773 0.50418517 0.50763969 0.51107732 0.51451308
 0.51790172 0.5212535  0.52459156 0.52790035 0.53118348 0.5344871
 0.53764993 0.54083183 0.54399565 0.54711016 0.55020864 0.55329183
 0.55632142 0.55933838 0.56231713 0.56528668 0.56823396 0.57115801
 0.57405632 0.57693449 0.5797856  0.58262309 0.58543629 0.58824039
 0.59104082 0.59378191 0.59651768 0.59923496 0.60192592 0.60460317
 0.6072485  0.60986488 0.61245087 0.61501759 0.61757333 0.62010882
 0.62262925 0.62512755 0.62761394 0.63006827 0.63251573 0.63491929
 0.63729676 0.63965505 0.64200451 0.64434361 0.64666582 0.6489463
 0.65121136 0.65346664 0.65571619 0.65795495 0.66017579 0.66238898
 0.66458656 0.66676602 0.66891151 0.67105236 0.6731873  0.67529454
 0.67737293 0.67942354 0.68145778 0.68347823 0.68548907 0.68749158
 0.68948048 0.69144798 0.6933975  0.69533843 0.6972612  0.69916593
...
```

MSM



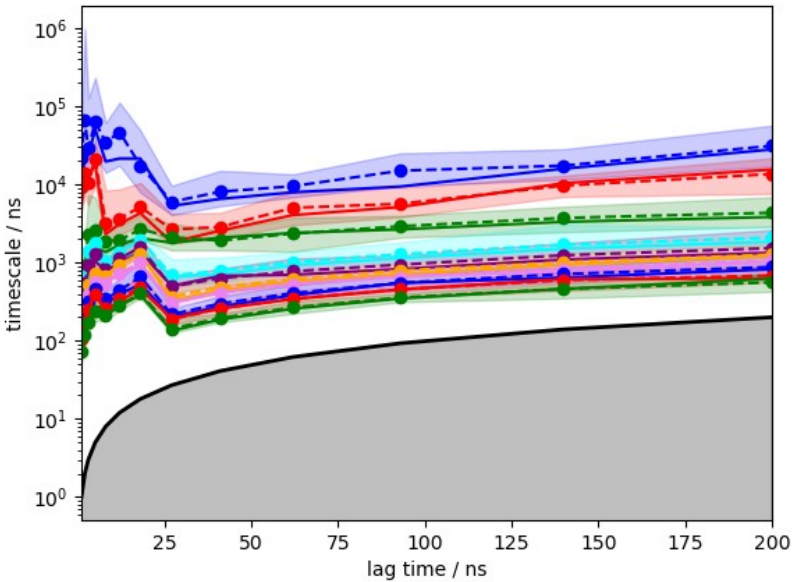
Better!



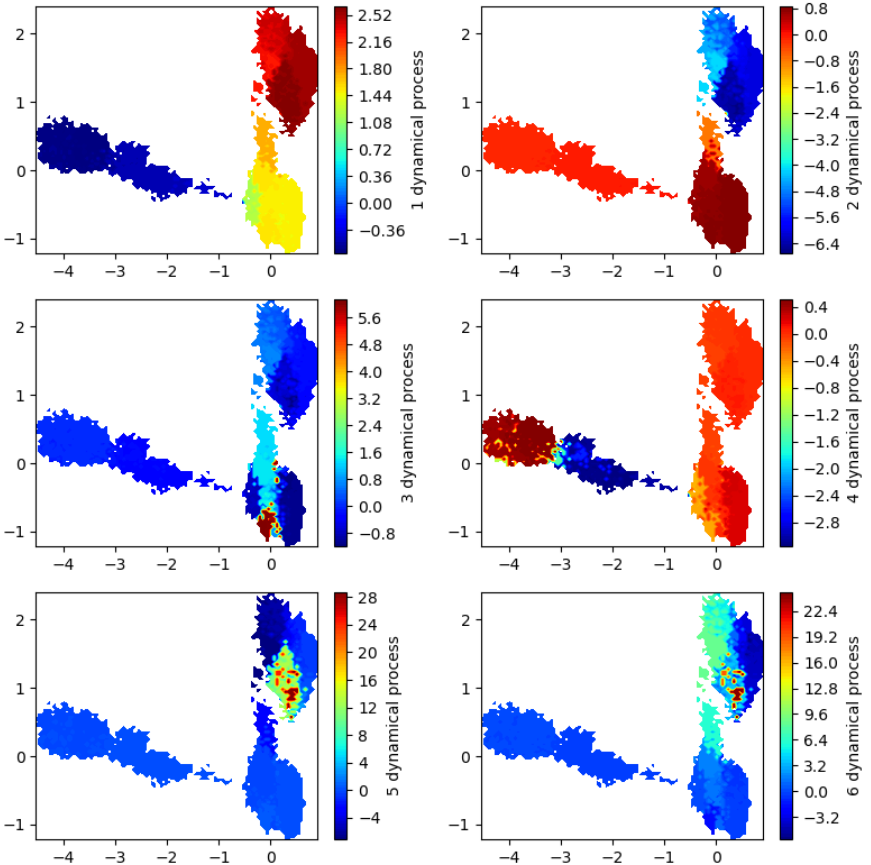
Processes (and macrostates) are not “clean”

MSM 2

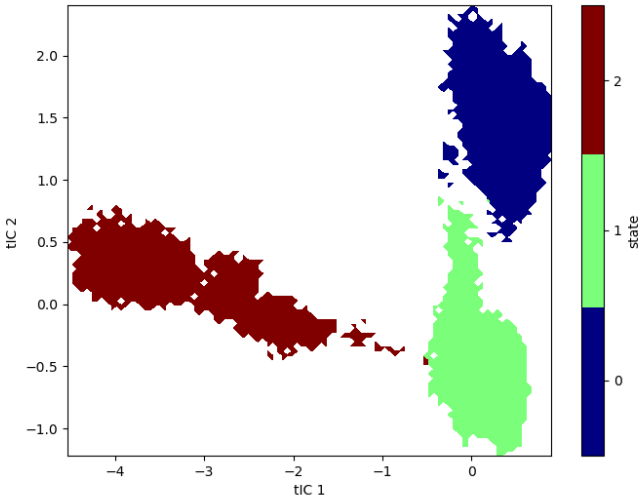
Cluster in 5 tICA dimensions instead of all dimensions giving 95% cumulative kinetic variance



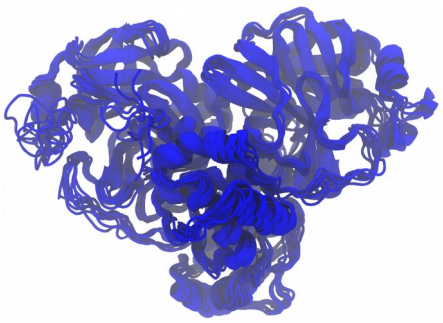
Timescales look weird…

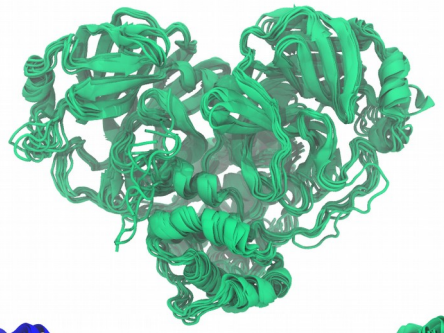


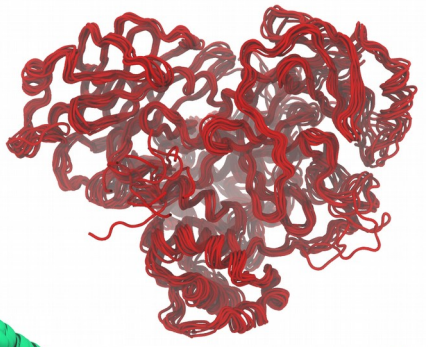
Macrostates

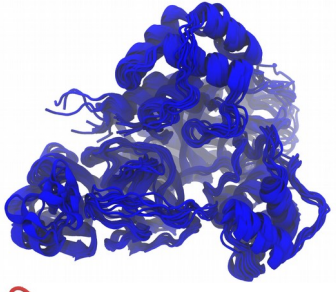


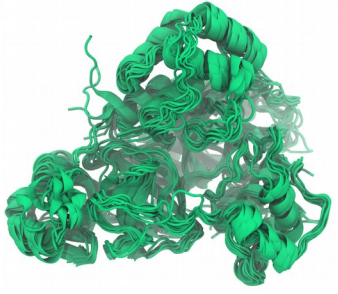
VMD couldn't understand secondary structure for some reason...

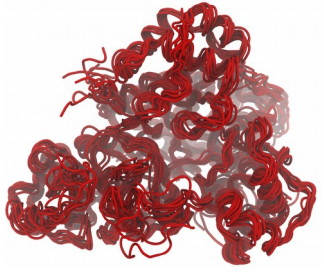


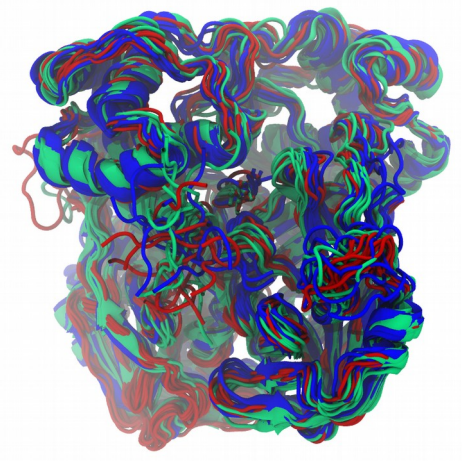


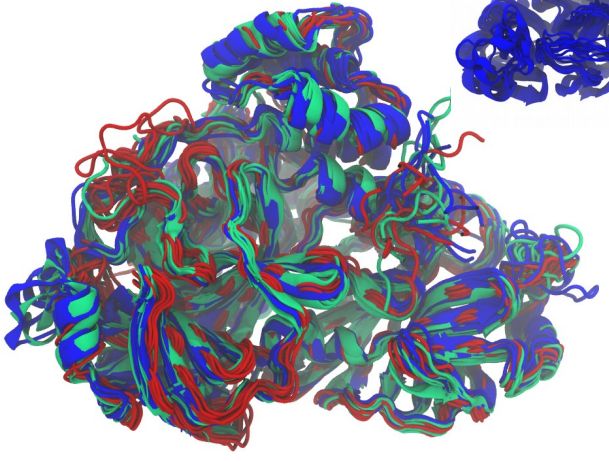






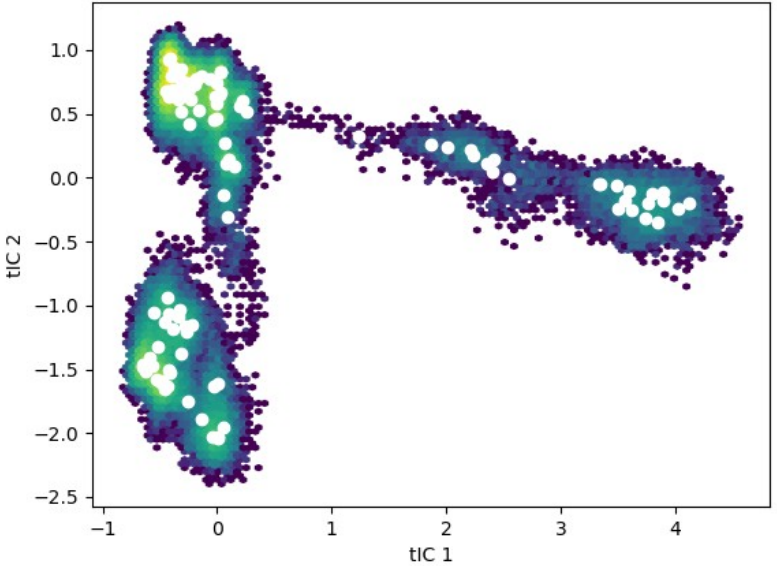






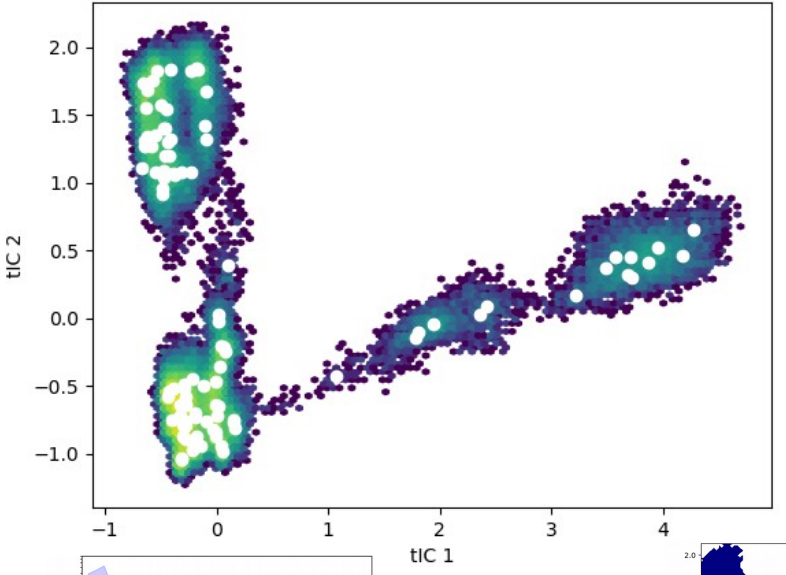
TICA (ENM-like distances)

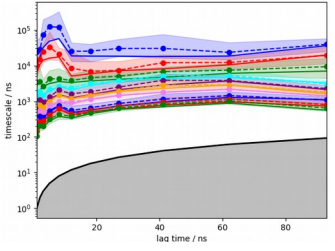
Cutoff_min: 4.8, cutoff_max: 5 (5889 distances)

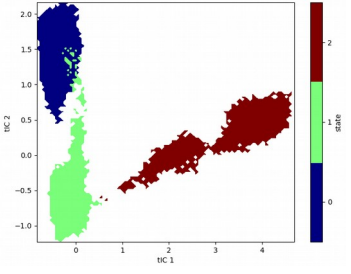


Similar as positions

Cutoff_min: 2.3, cutoff_max: 2.5 (9102 distances)

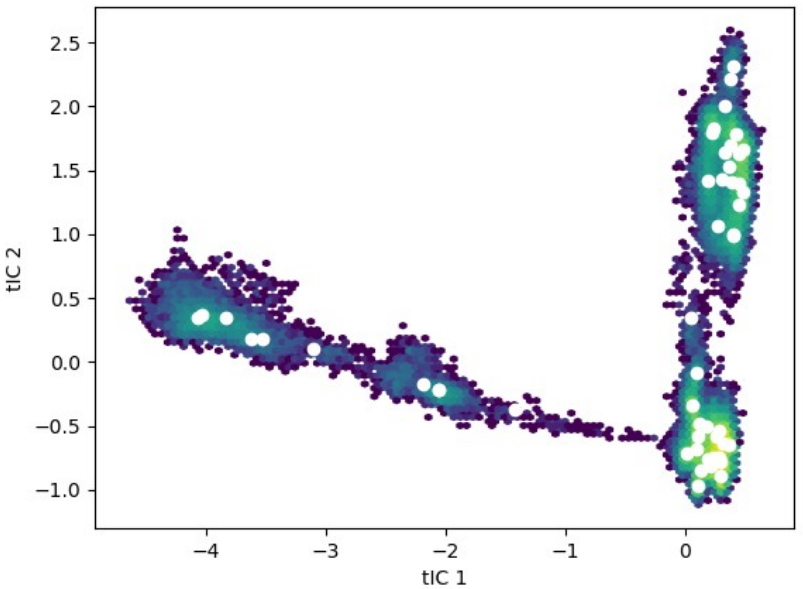


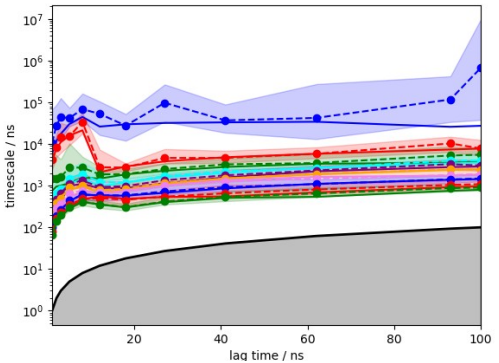




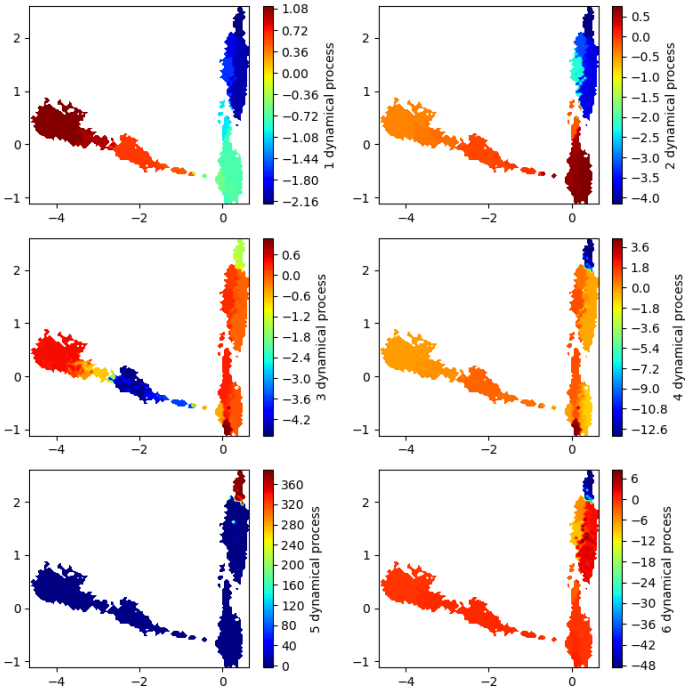
TICA (ENM-like distances)

Cutoff_min: 0.5, cutoff_max: 1.3 (10668 distances)

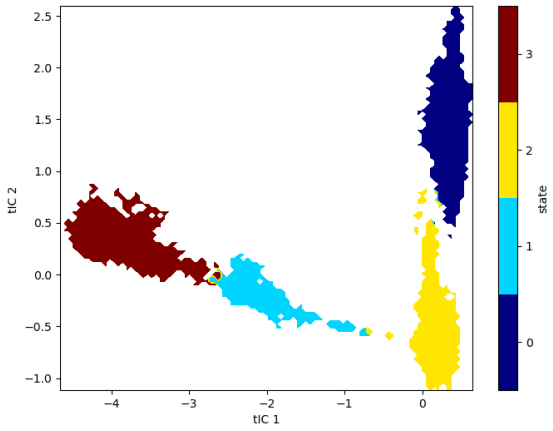


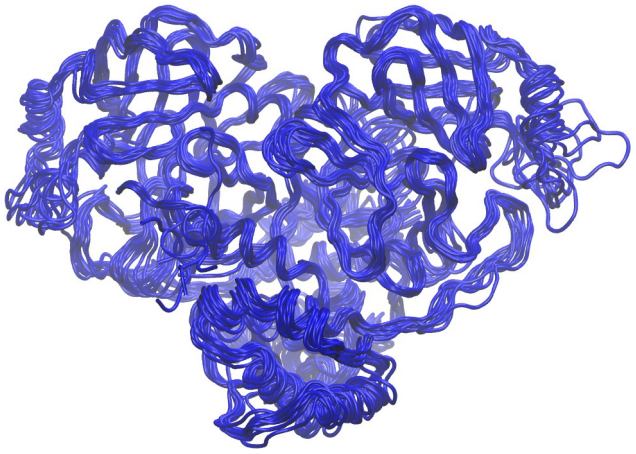


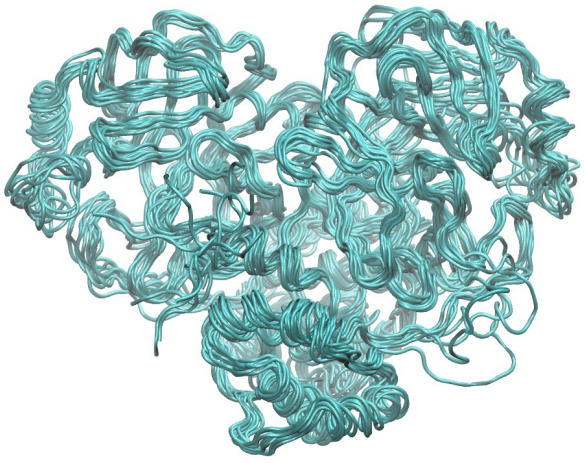
Note: units are not necessarily ns, but rather the time between two frames in the simulation, sorry!

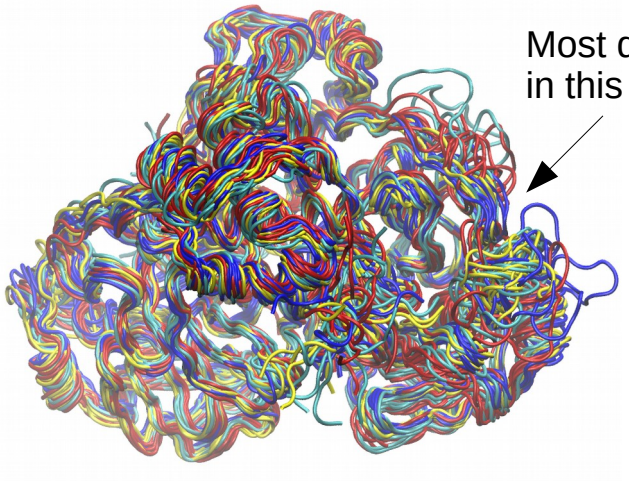


Macrostates









Most differences
in this part