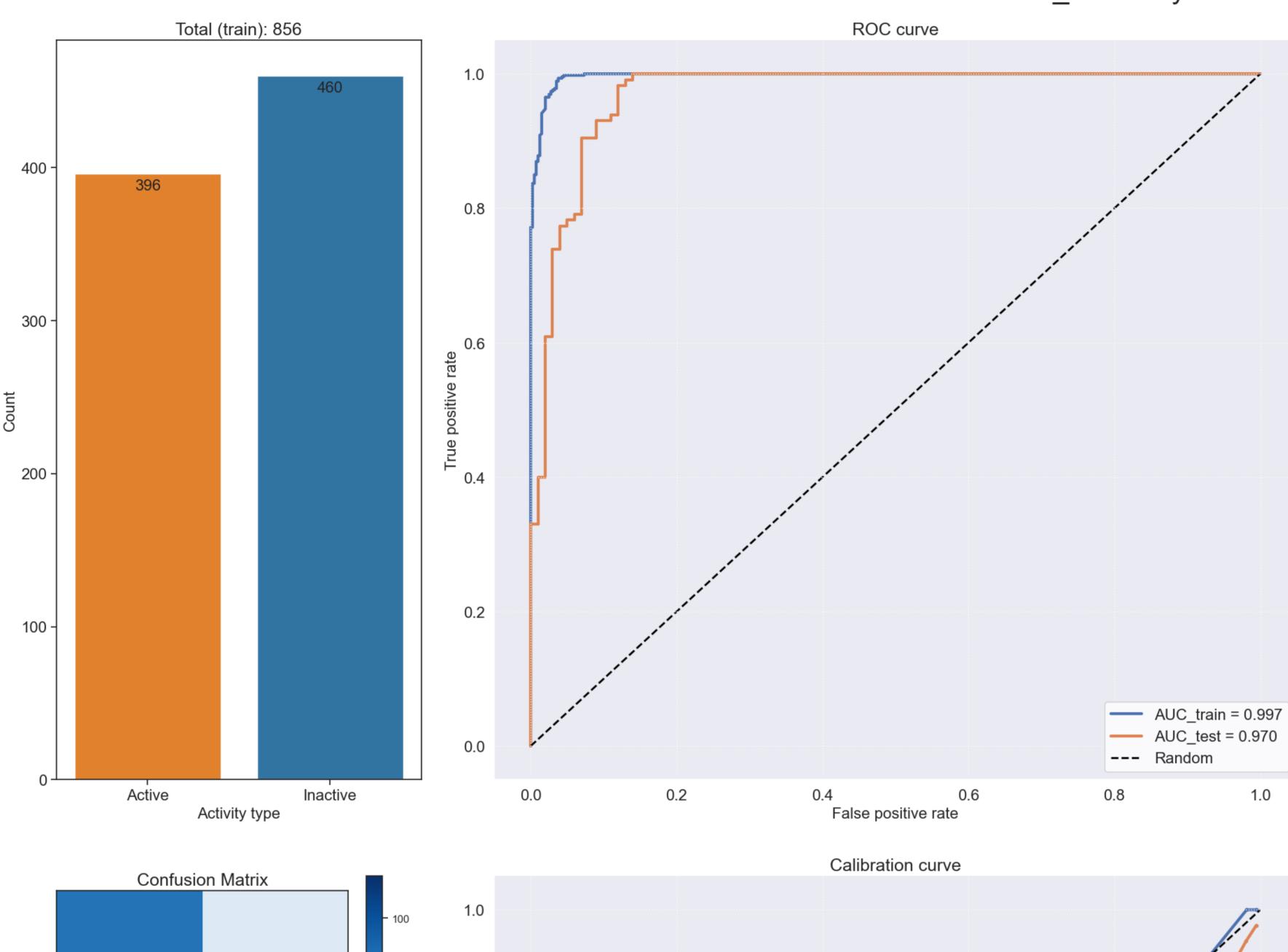
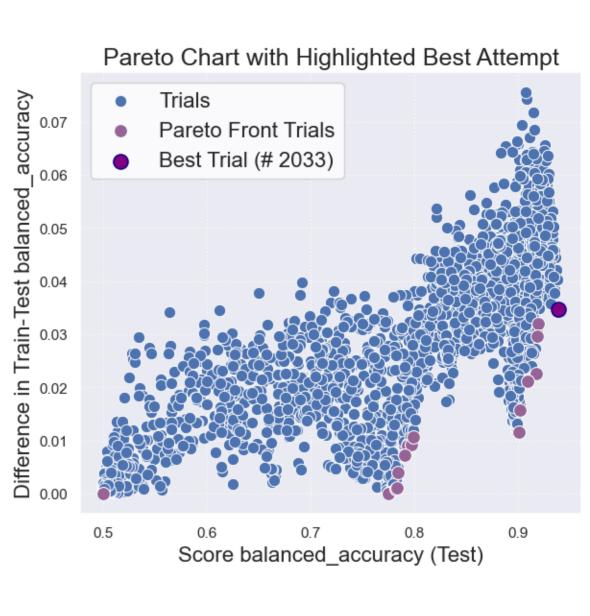
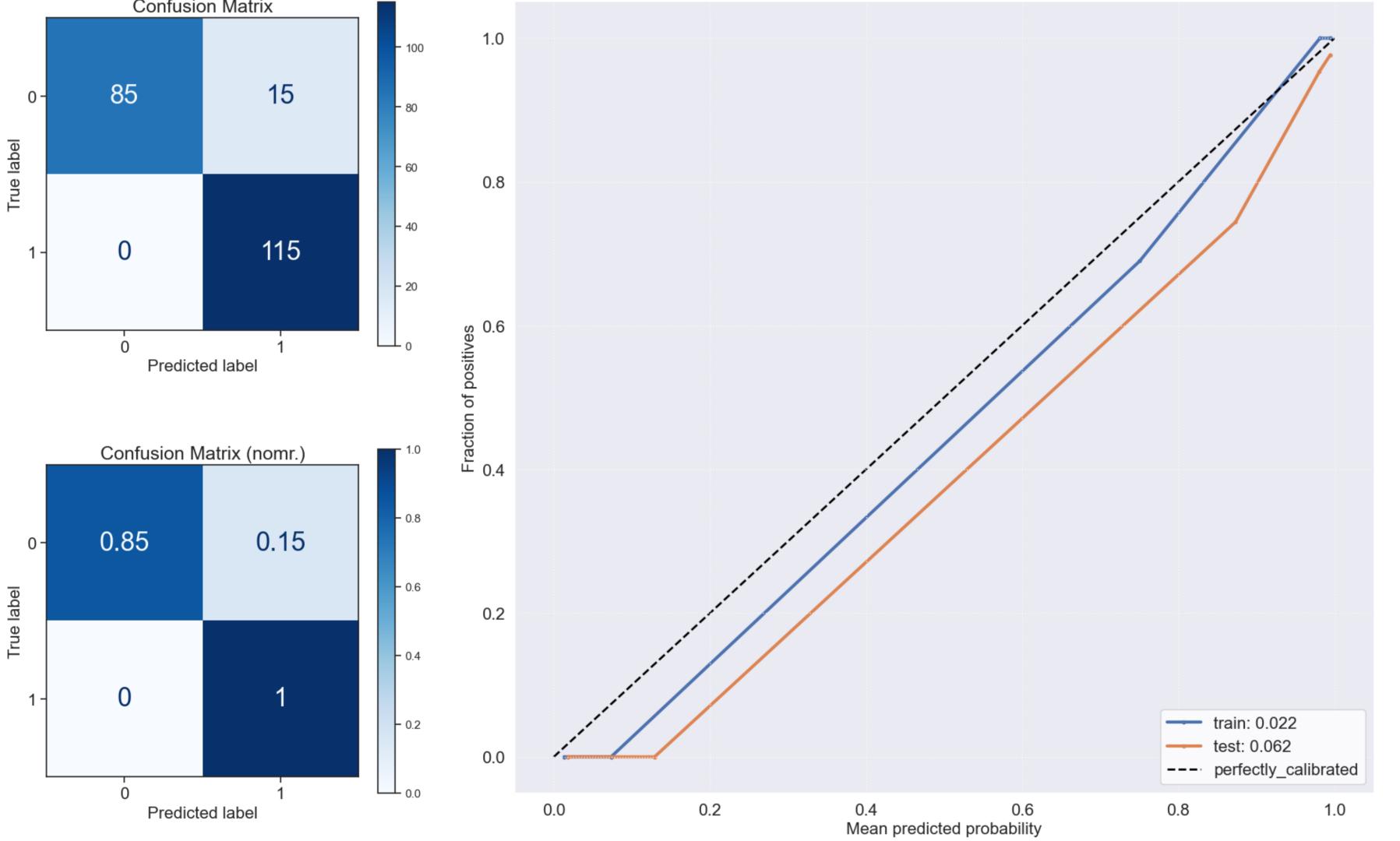
## Protein: P31639 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	4.6957
colsample_bytree	0.6118
gamma	0.7692
lambda	4.2774
learning_rate	0.0581
max_depth	9.0
min_child_weight	1.0
n_estimators	563.0
reg_alpha	0.8197
reg_lambda	0.735
scale_pos_weight	2.2383
subsample	0.6699

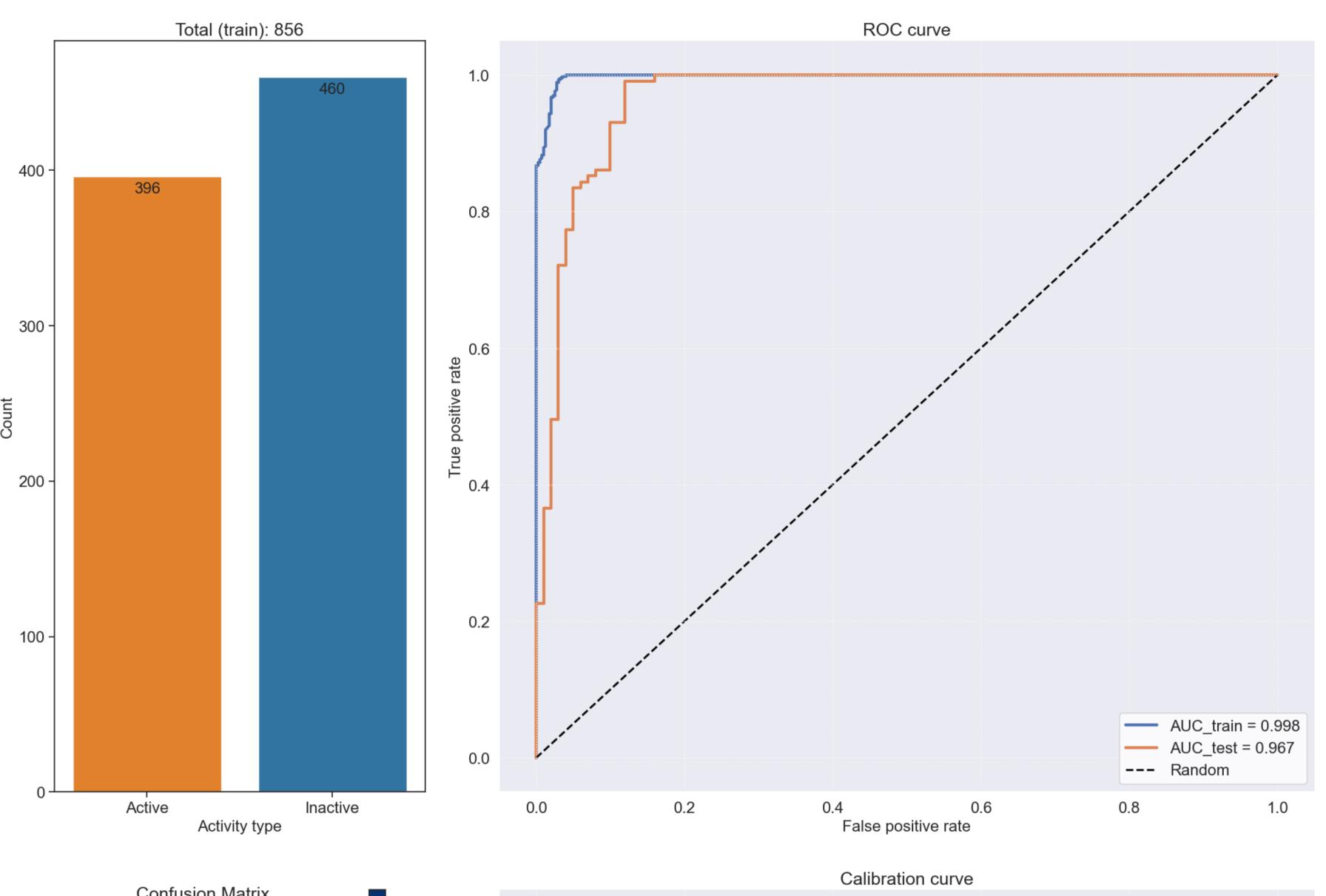


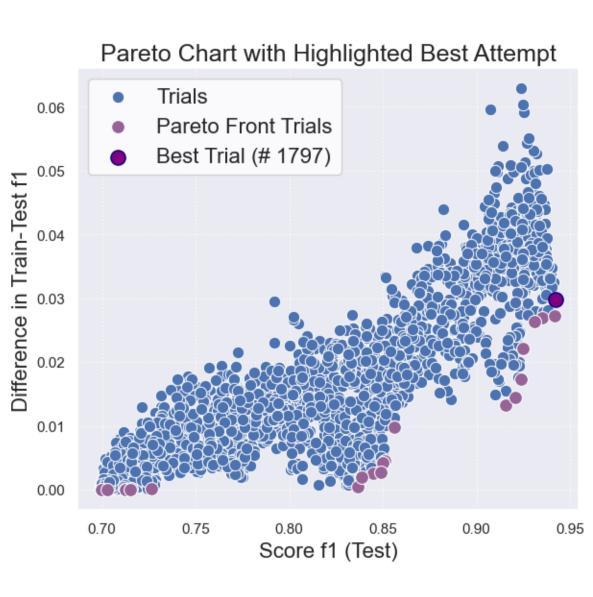
25	lna					zone a	ecoy (215)
00							
25							
20							
15							
5						_ [	
0							10
(	0.0	.∠	0.4 Proba	0.6 ability	U	.ŏ	1.0

Active/Inactive probability

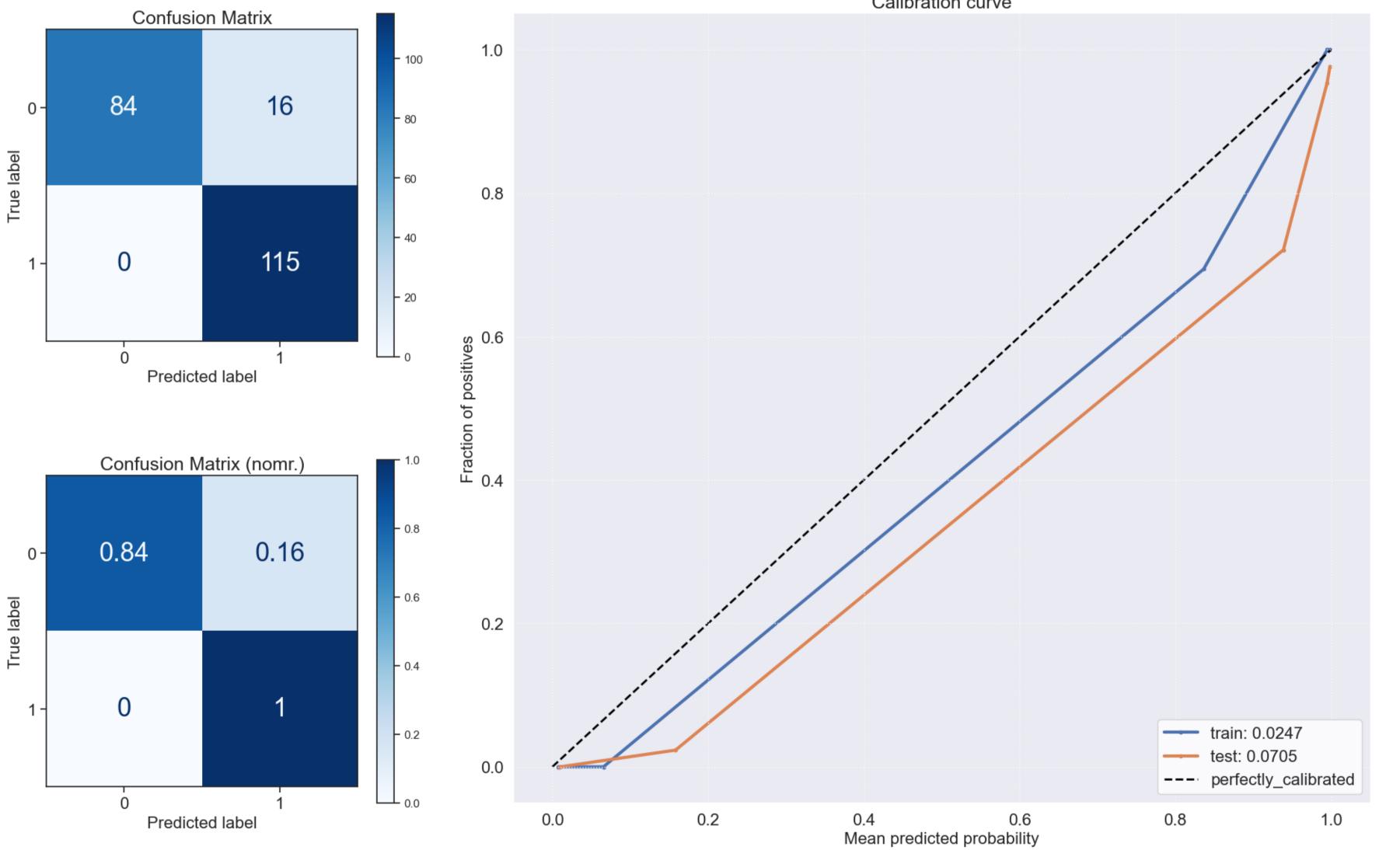
Metric	Train	Test
AUC	0.9969	0.9701
accuracy	0.9778	0.9302
balanced accuracy	0.9762	0.925
recall_(sens)	0.9978	1.0
specificity	0.9545	0.85
precision	0.9623	0.8846
f1_score	0.9778	0.9295

## Protein: P31639 - metric: f1





XGBoost Parameters	Value
alpha	2.8385
colsample_bytree	0.8336
gamma	0.6683
lambda	1.7017
learning_rate	0.0373
max_depth	9.0
min_child_weight	3.0
n_estimators	496.0
reg_alpha	0.1226
reg_lambda	0.0022
scale_pos_weight	7.5107
subsample	0.9116

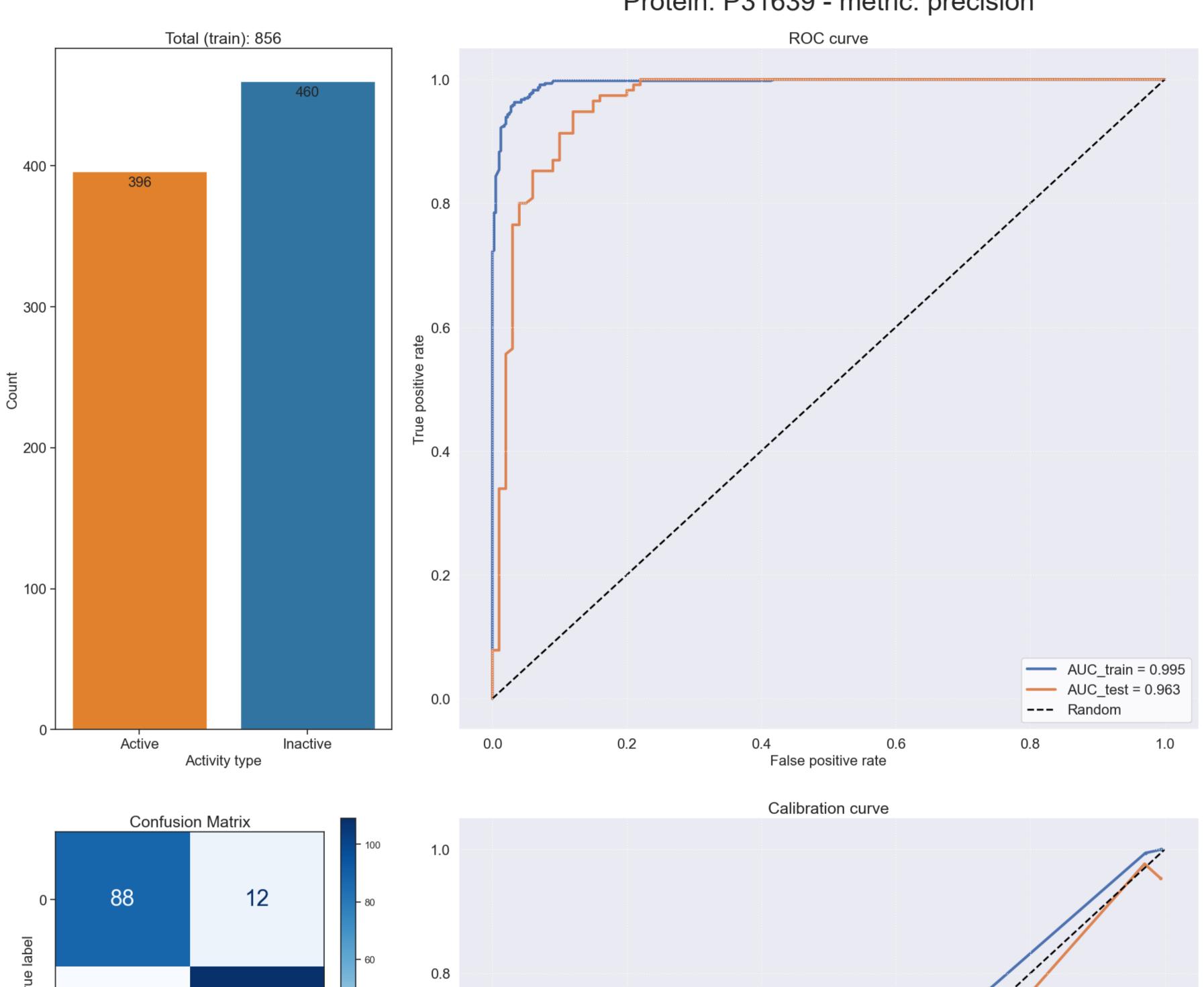


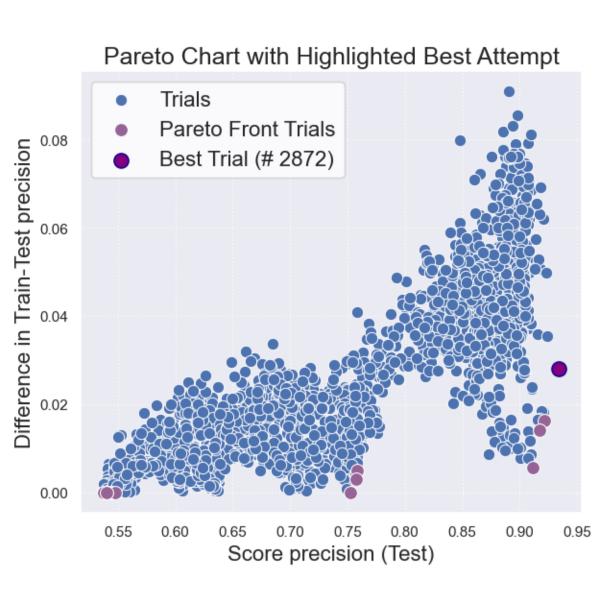
70		ctive zone		Active	zone active inacti	e (115) ve (100)
60					deco	y (215)
50						
Percent						
₾ 30						
20						
10				_		
0 —	0.0 0	.2 0.4	Probability	0.6	.8	1.0

Active/Inactive probability

Metric	Train	Test
AUC	0.9977	0.9666
accuracy	0.9696	0.9256
balanced accuracy	0.9672	0.92
recall_(sens)	1.0	1.0
specificity	0.9343	0.84
precision	0.9465	0.8779
f1_score	0.9695	0.9248

## Protein: P31639 - metric: precision



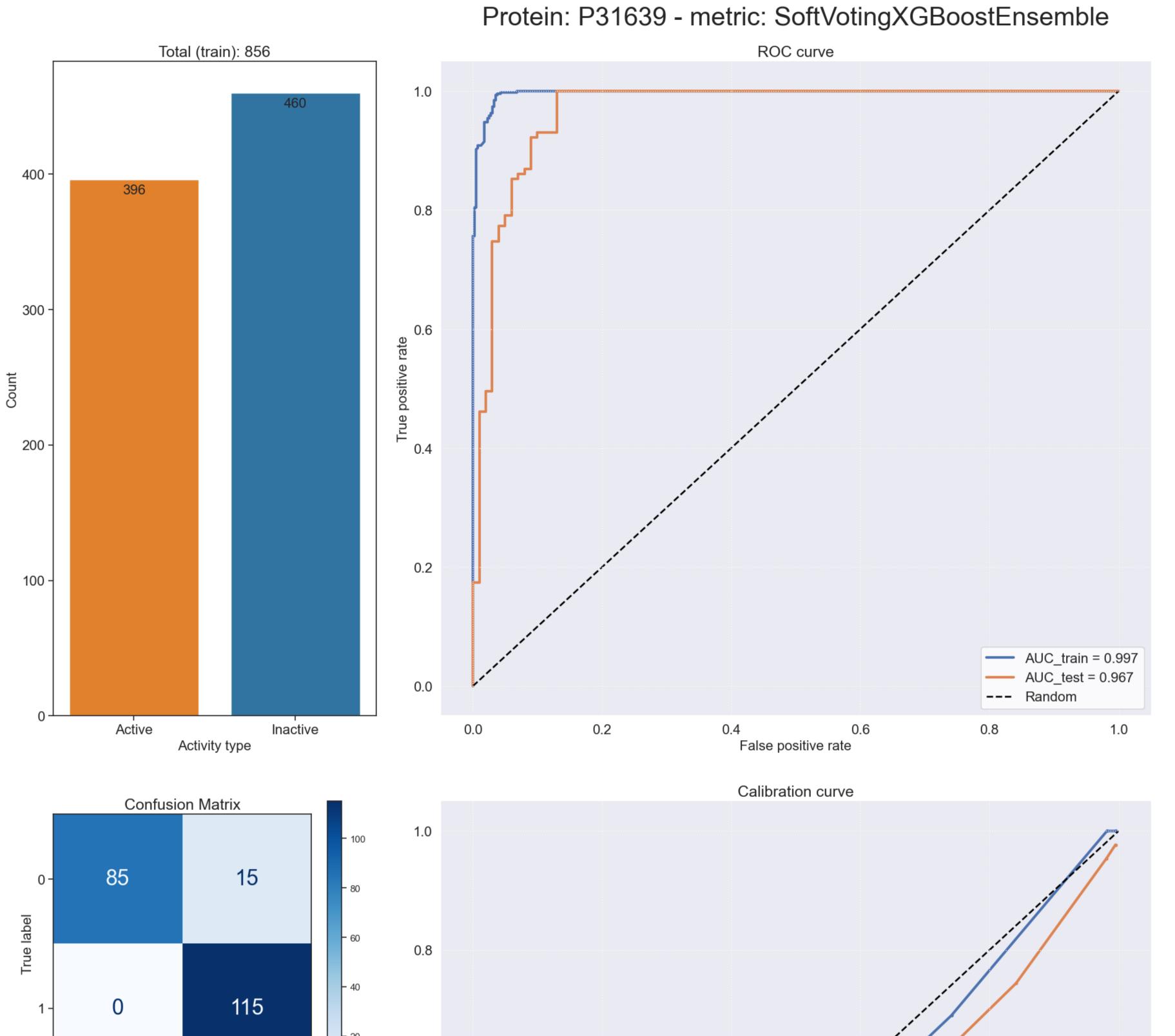


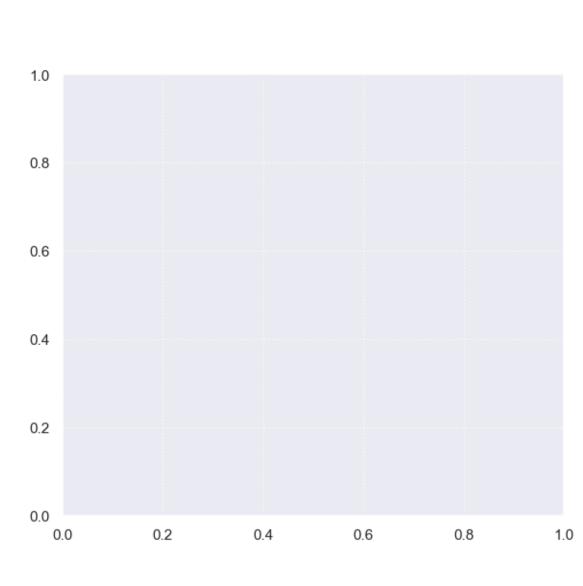
XGBoost Parameters	Value
alpha	4.2006
colsample_bytree	0.9426
gamma	0.0211
lambda	3.6519
learning_rate	0.1715
max_depth	9.0
min_child_weight	4.0
n_estimators	300.0
reg_alpha	0.5501
reg_lambda	0.7094
scale_pos_weight	1.0642
subsample	0.5573

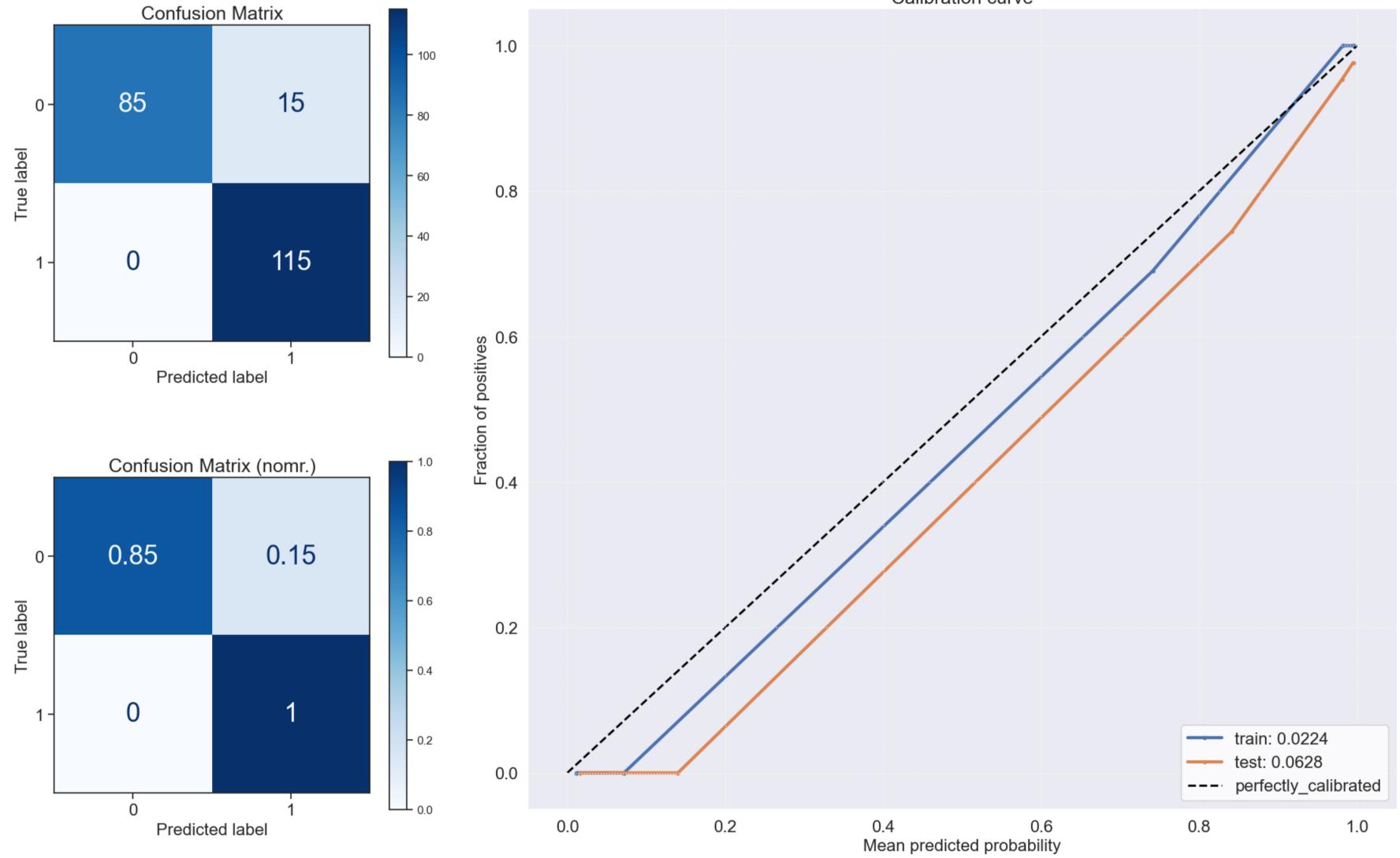
0 -	88	12	- 100 - 80	1.0		
True label	6	109	- 60 - 40	0.8		
	0 Predicte	1 ed label	- 20	n of positives		
	Confusion M	latrix (nomr.)	- 0.9	Fraction 6.0		
abel O	0.88	0.12	- 0.8 - 0.7 - 0.6	0.2		
True label	0.052	0.95	- 0.5 - 0.4 - 0.3 - 0.2 - 0.1	0.0		test: 0.0693
	0 Predicte	1 ed label		C	0 0.2 0.4 0.6 0.8 Mean predicted probability	1.0

25	Active/Inactive probability						
35		active zone		Activ	ve zone	type l active (115) l inactive (100)	
						decoy (215)	
25							
Percent							
10							
- 10							
5							
0	0.0 0	0.2		0.6	0.8	1.0	
i	Probability						

Metric	Train	Test
AUC	0.9946	0.9628
accuracy	0.965	0.9163
balanced accuracy	0.9651	0.9139
recall_(sens)	0.963	0.9478
specificity	0.9672	0.88
precision	0.9715	0.9008
f1_score	0.965	0.916







40	<u>-</u>						type	
	Ina	ctive zone			Active	zone	active (115 inactive (1	-
35							decoy (21	5)
20								
30								
± 25								
Percent 05 C								
о 20 <del></del>								
15								
10								
5								
0		<b>—</b> <u>C</u>						
U	0.0	2	0.4	0.6	(	0.8	1.0	
	Probability							

Active/Inactive probability

Metric	Train	Test
AUC	0.997	0.9674
accuracy	0.9778	0.9302
balanced accuracy	0.9762	0.925
recall_(sens)	0.9978	1.0
specificity	0.9545	0.85
precision	0.9623	0.8846
f1_score	0.9778	0.9295