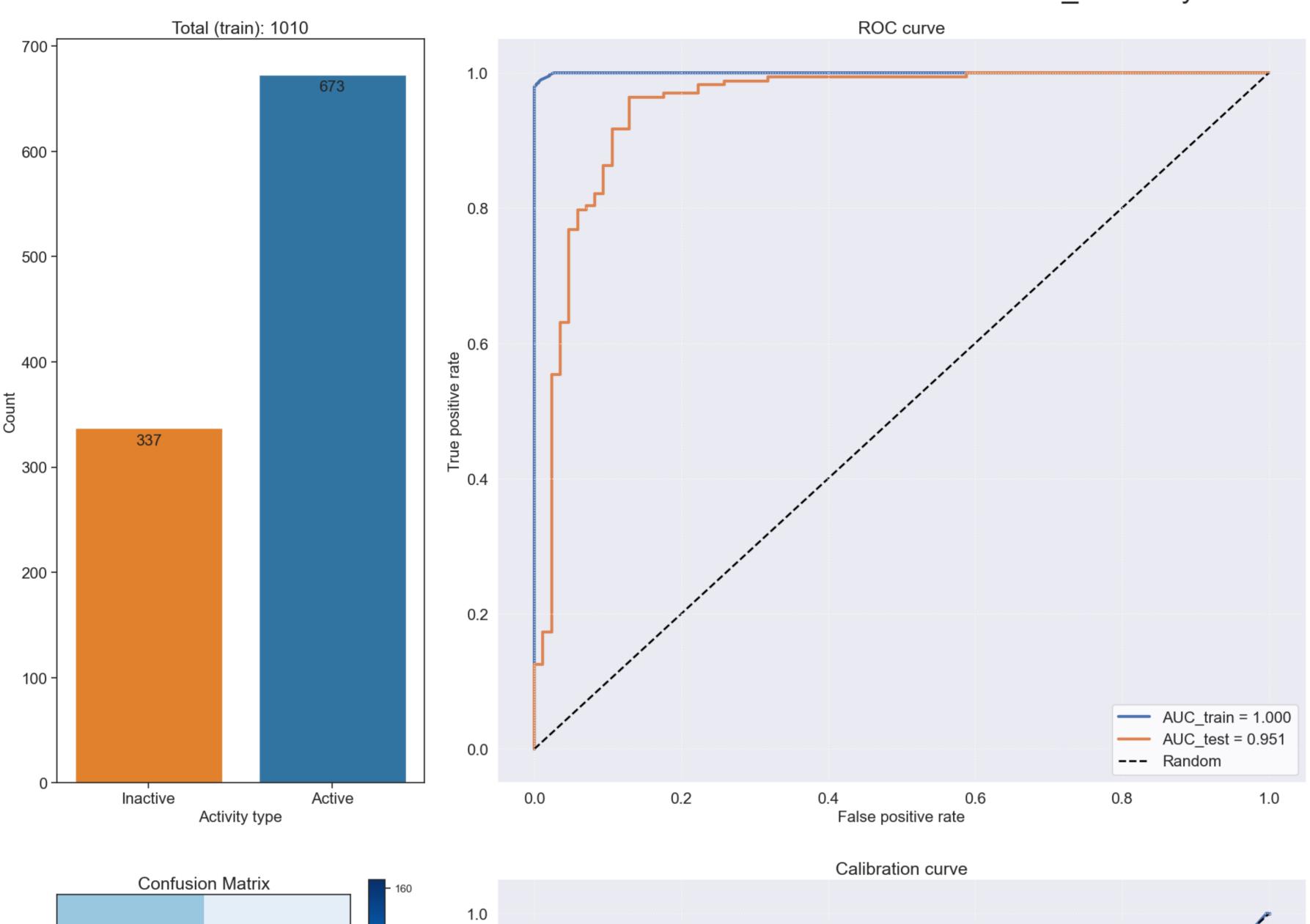
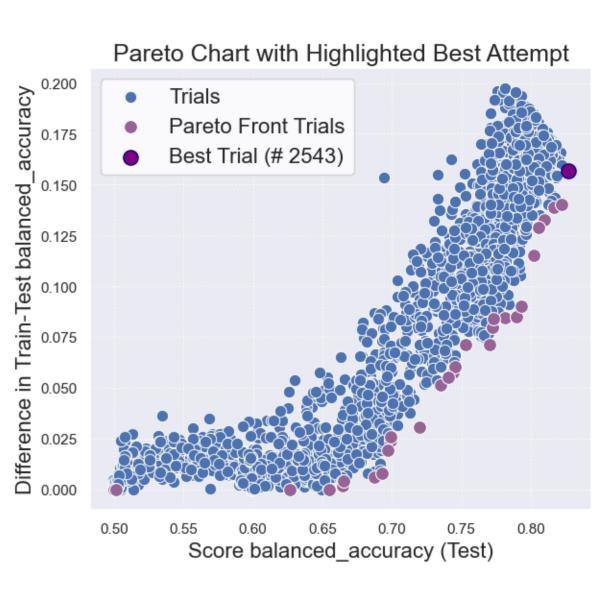
## Protein: O00519 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	4.9439
colsample_bytree	0.8633
gamma	0.2471
lambda	3.1302
learning_rate	0.8504
max_depth	10.0
min_child_weight	1.0
n_estimators	129.0
reg_alpha	0.7933
reg_lambda	0.6883
scale_pos_weight	8.8991
subsample	0.8275

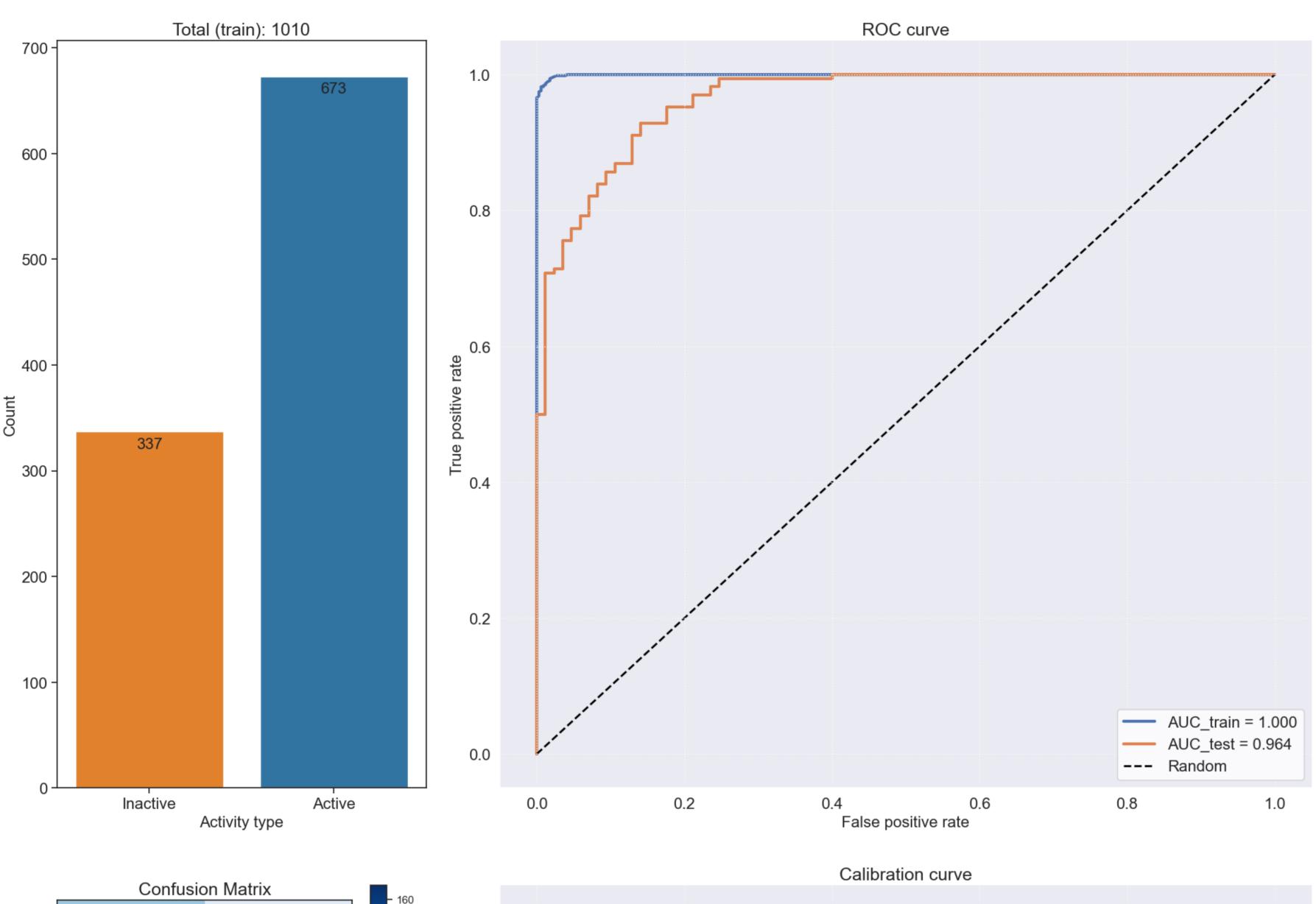
0 -	66	19	- 120							
True label			- 100	0.0						
True			- 80	0.8						
1 -	4	164	- 60							
			- 40							
	0 Predicte	1 ed label	20	of positives			,,,			
	Confusion M	latrix (nomr.)		Fraction 6.0			parara			
bel -	0.78	0.22	- 0.8 - 0.6			, or				
True label	0.024	0.98	- 0.4 - 0.2	0.2					train: 0.0 test: 0.07	
	0 Predicte	1 ed label			0.0	0.2	0.4	0.6	0.8	1.0

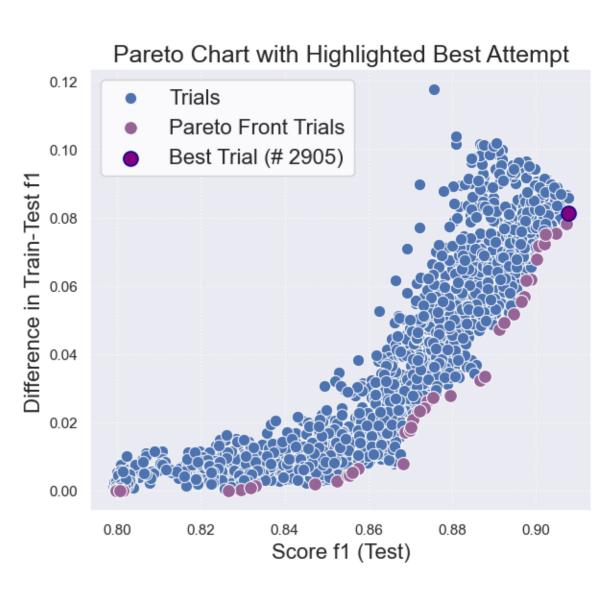
Mean predicted probability

	Ac	tive/Inactive probability	
80	Inactive zone		Active zone type active (168) inactive (85)
70			
60			
ent 50			
Percent 0 6			
30			
20			
0 —			
Ü	0.0 0.2	0.4 0.6 Probability	0.8 1.0

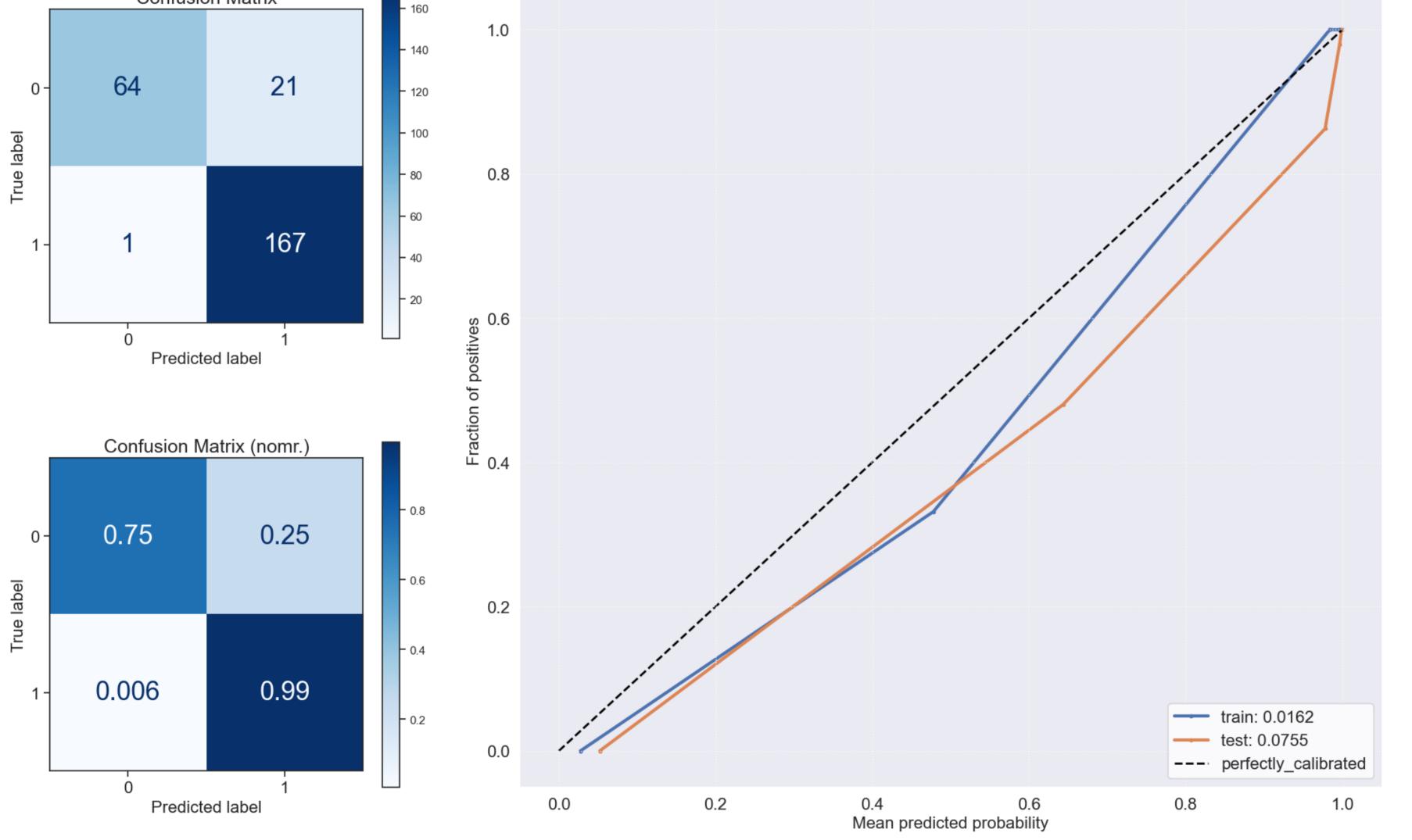
Metric	Train	Test
AUC	0.9998	0.9513
accuracy	0.9901	0.9091
balanced accuracy	0.9852	0.8763
recall_(sens)	1.0	0.9762
specificity	0.9703	0.7765
precision	0.9854	0.8962
f1_score	0.9901	0.9066

## Protein: O00519 - metric: f1





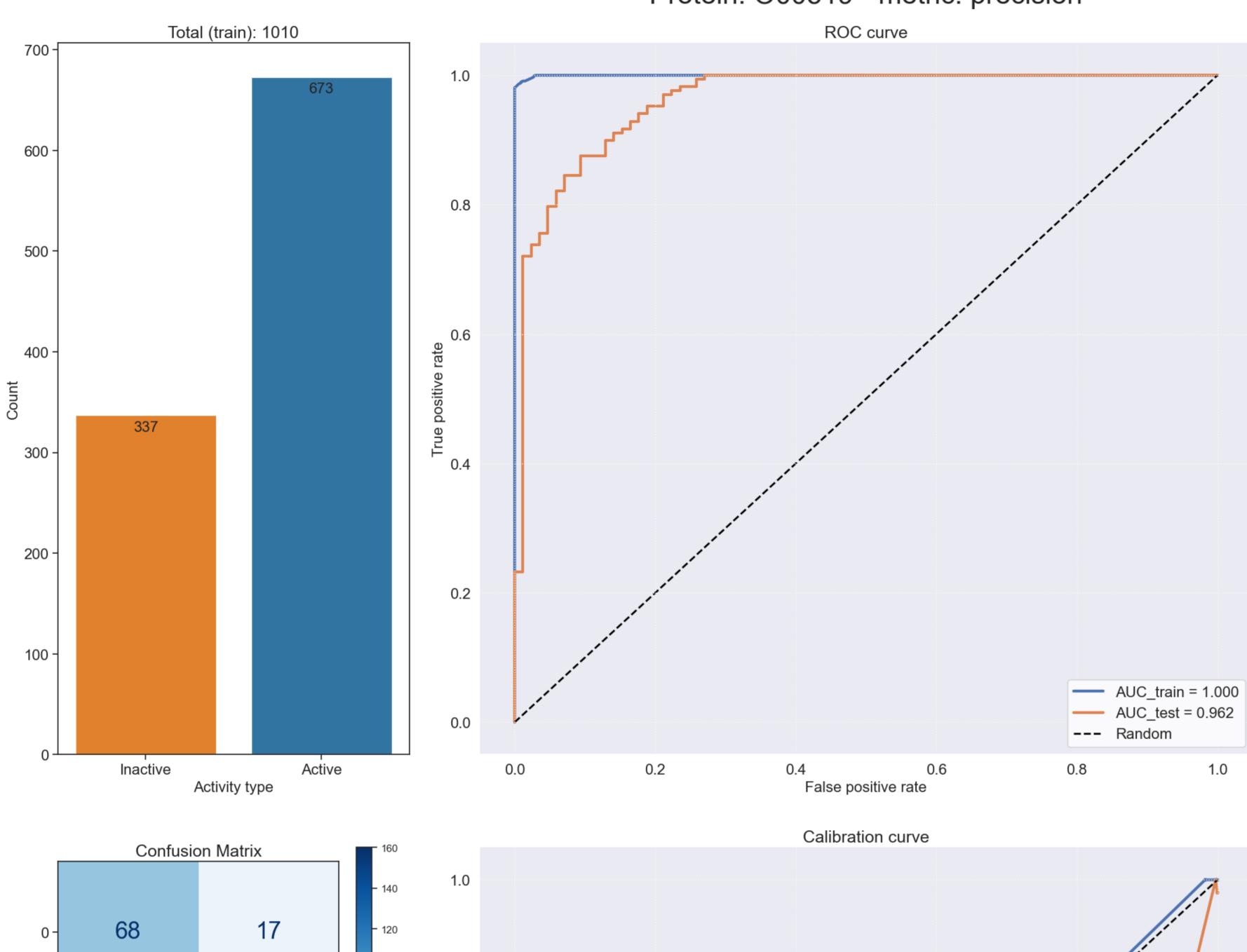
XGBoost Parameters	Value
alpha	4.8129
colsample_bytree	0.7838
gamma	0.0075
lambda	2.1785
learning_rate	0.0532
max_depth	7.0
min_child_weight	1.0
n_estimators	580.0
reg_alpha	0.801
reg_lambda	0.1089
scale_pos_weight	3.9168
subsample	0.7022

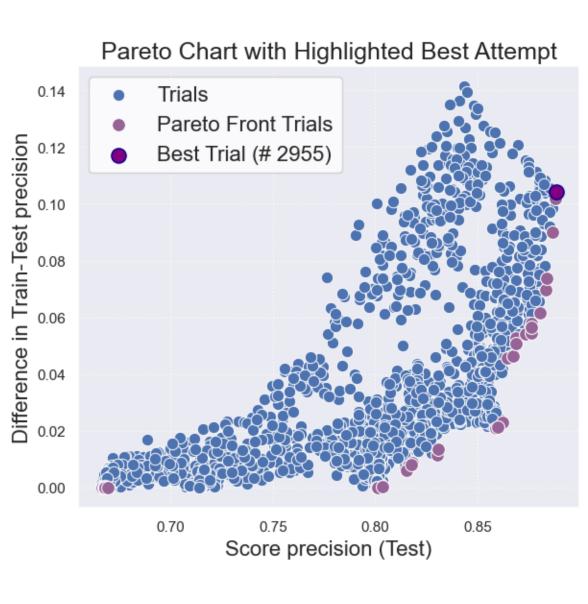


		Active/Inactiv	e probability		
70	Inactive zone			Active zone	active (168) inactive (85)
50					decoy (253)
Percent					
g 30					
20					
0					
	0.0 0.2	0.4 Proba	0.6 bility	0.8	1.0

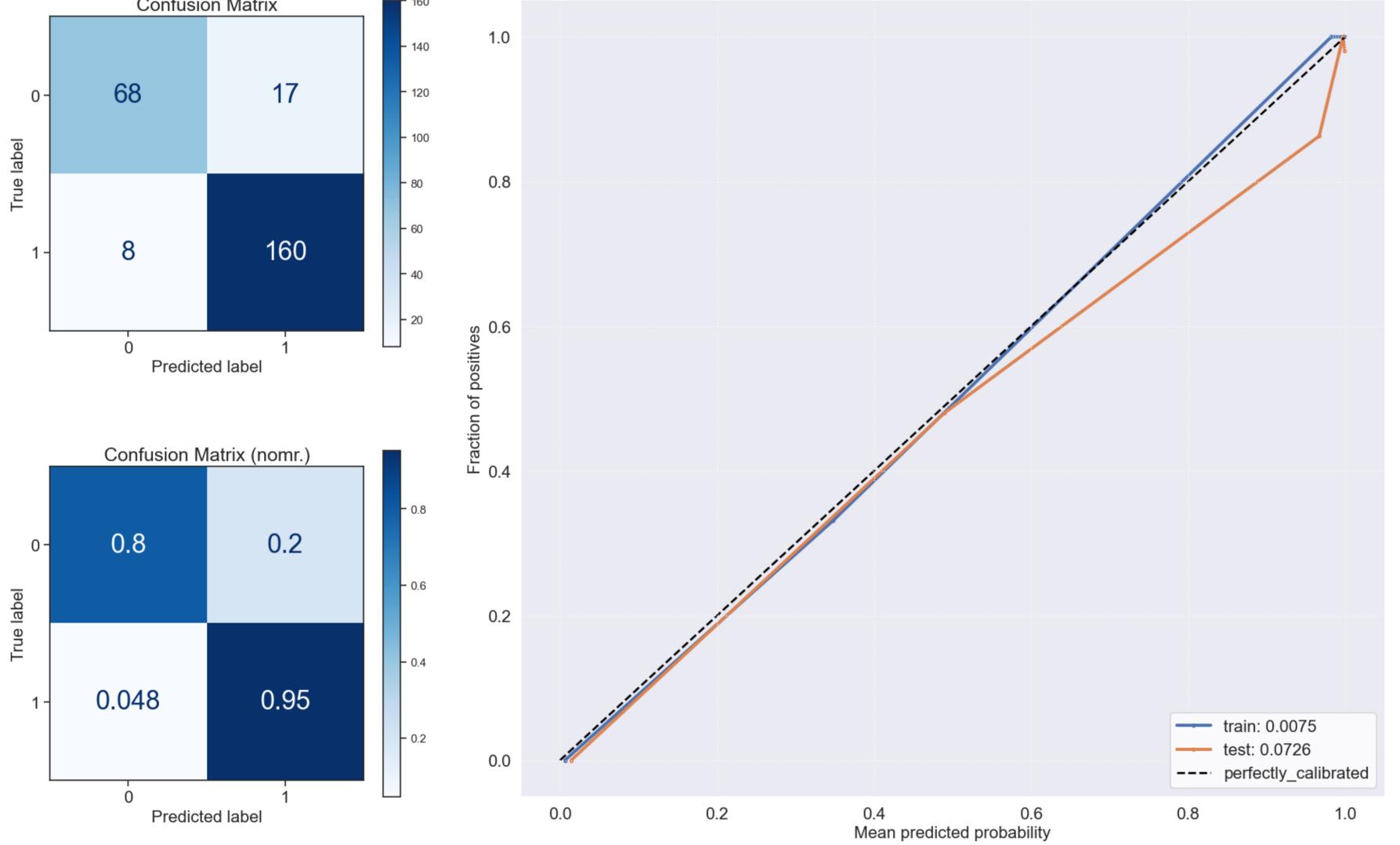
Metric	Train	Test
AUC	0.9996	0.9635
accuracy	0.9822	0.913
balanced accuracy	0.9733	0.8735
recall_(sens)	1.0	0.994
specificity	0.9466	0.7529
precision	0.974	0.8883
f1_score	0.9821	0.9097

## Protein: O00519 - metric: precision





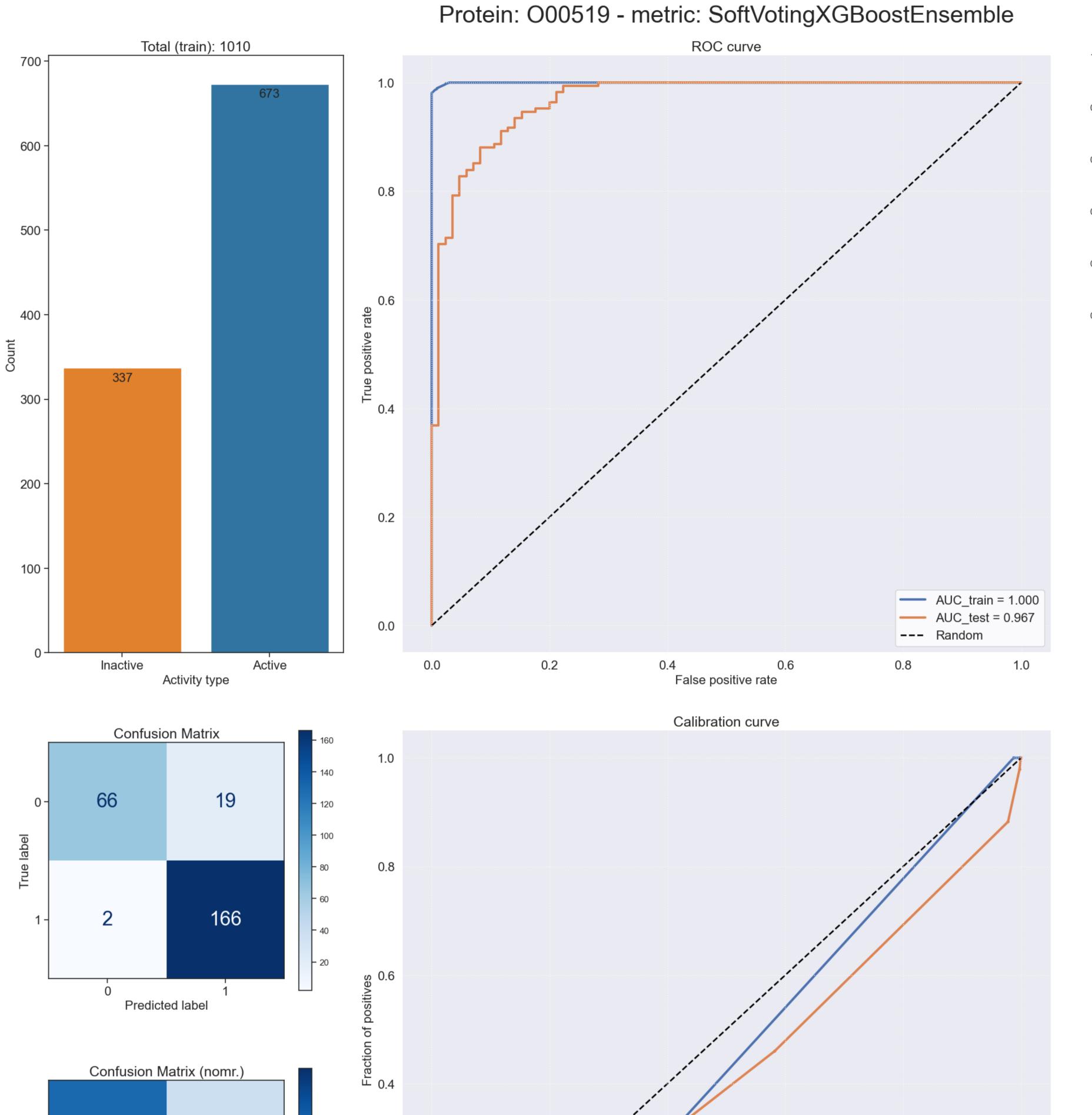
XGBoost Parameters	Value
alpha	4.9127
colsample_bytree	0.8088
gamma	0.1147
lambda	4.8615
learning_rate	0.3734
max_depth	7.0
min_child_weight	1.0
n_estimators	921.0
reg_alpha	0.0118
reg_lambda	0.5397
scale_pos_weight	1.0463
subsample	0.9553

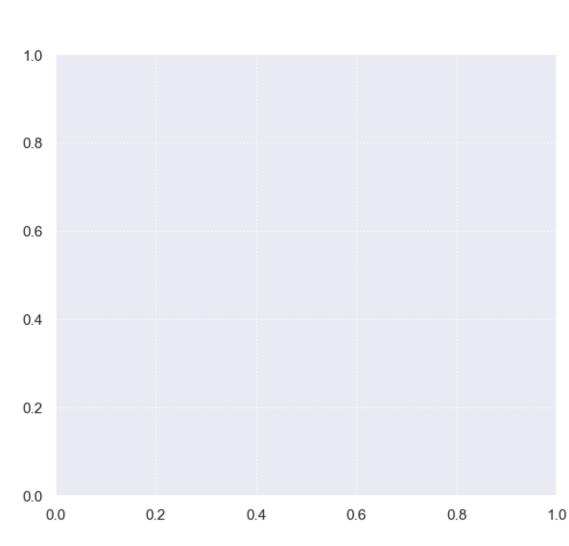


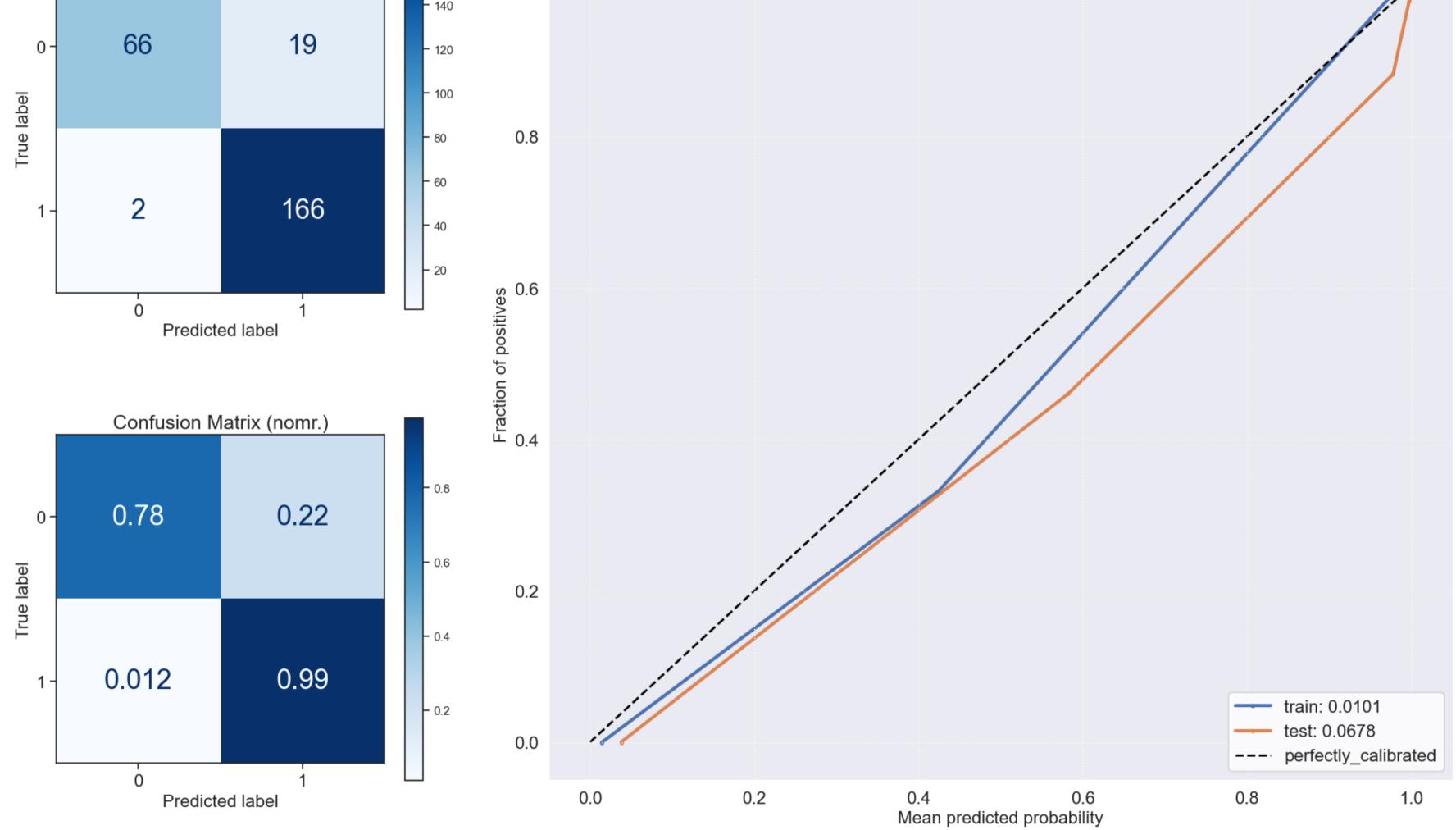
00	lna	active zone			Active zone	type active (168)
50						decoy (253)
Percent						
20						
0 —	0.0 0	.2	0.4	0.6	0.8	1.0
			Probab			

Active/Inactive probability

Metric	Train	Test
AUC	0.9997	0.9622
accuracy	0.9891	0.9012
balanced accuracy	0.9874	0.8762
recall_(sens)	0.9926	0.9524
specificity	0.9822	0.8
precision	0.9911	0.904
f1_score	0.9891	0.8997







	Ac	tive/Inactive probab	ility		
70 -·····	Inactive zone		Active	inacti	e (168) ve (85)
50				decoy	(253)
eut 40					
Percent 30					
20					
10					
0-	0.0 0.2 0.	.4 0 Probability	.6 0	).8 1	.0

Metric	Train	Test
AUC	0.9998	0.9673
accuracy	0.9901	0.917
balanced accuracy	0.9852	0.8823
recall_(sens)	1.0	0.9881
specificity	0.9703	0.7765
precision	0.9854	0.8973
f1_score	0.9901	0.9144