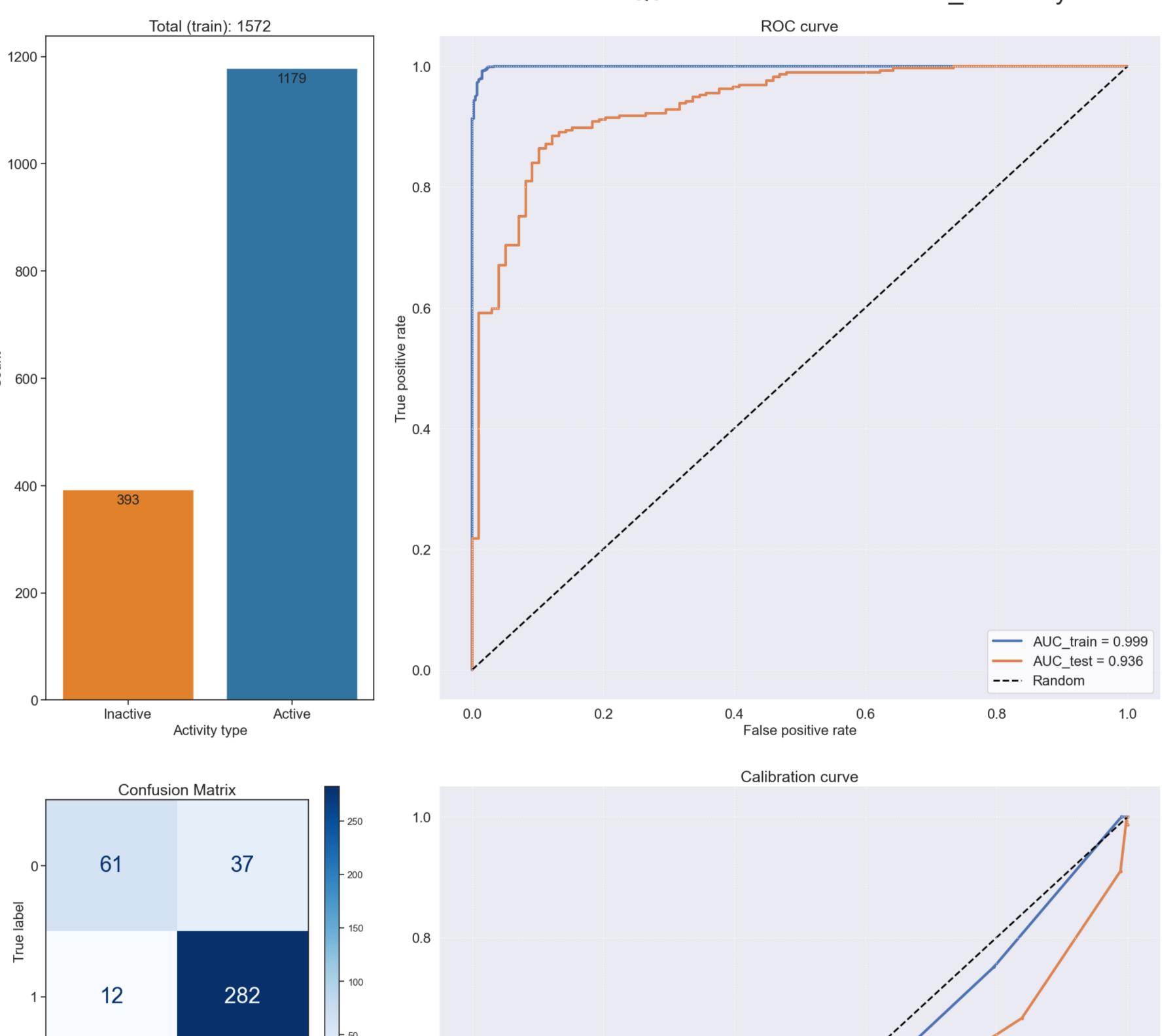
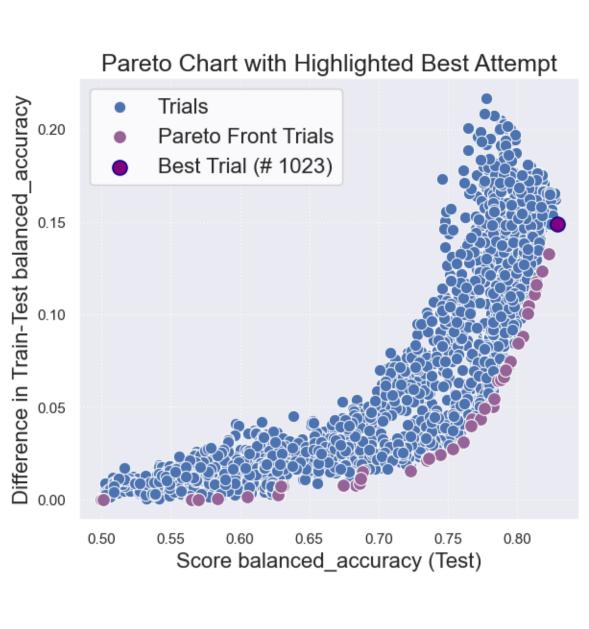
Protein: Q8NER1 - metric: balanced_accuracy





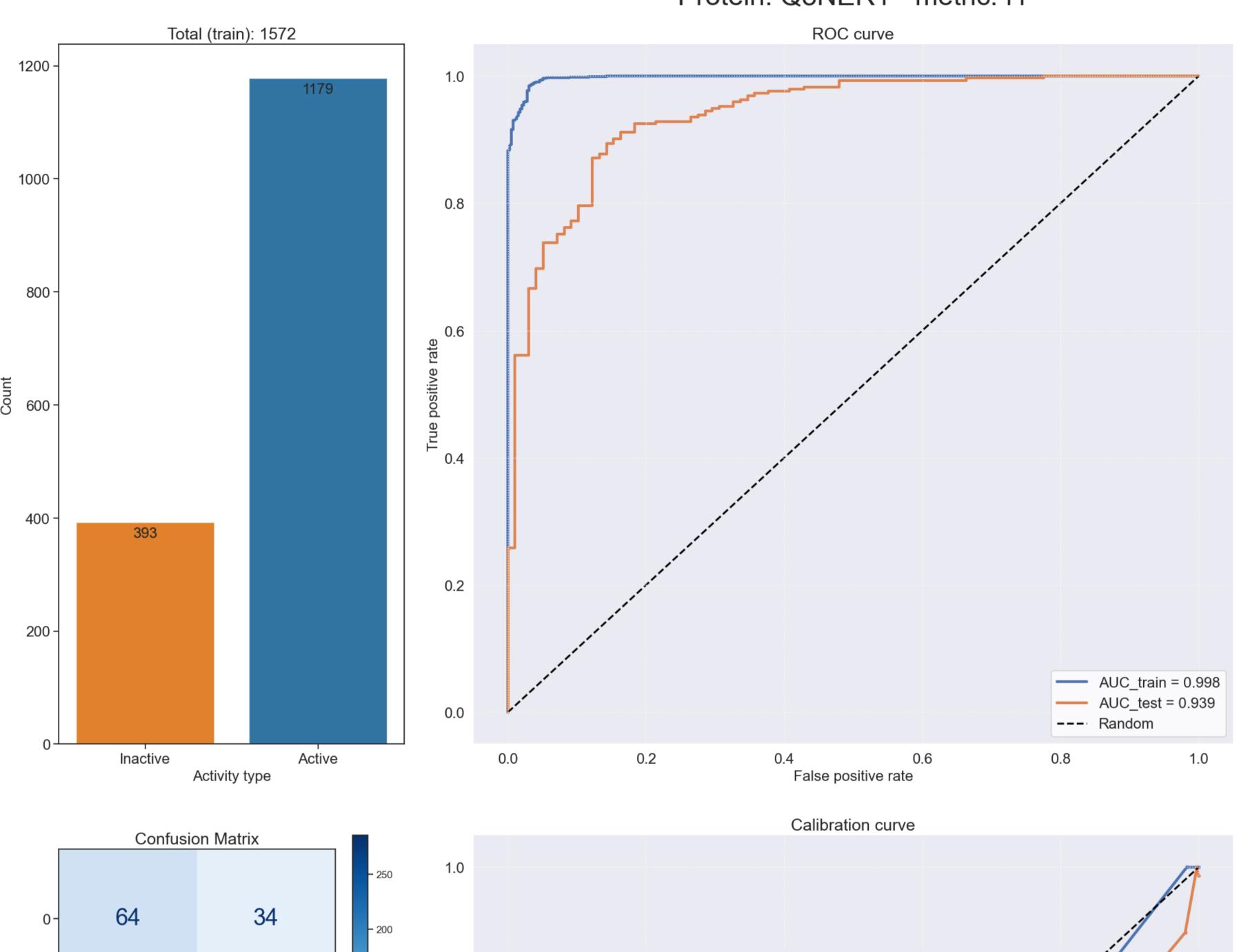
XGBoost Parameters	Value
alpha	1.2209
colsample_bytree	0.9095
gamma	0.42
lambda	0.2793
learning_rate	0.1701
max_depth	3.0
min_child_weight	2.0
n_estimators	916.0
reg_alpha	0.0389
reg_lambda	0.9192
scale_pos_weight	2.3338
subsample	0.6844

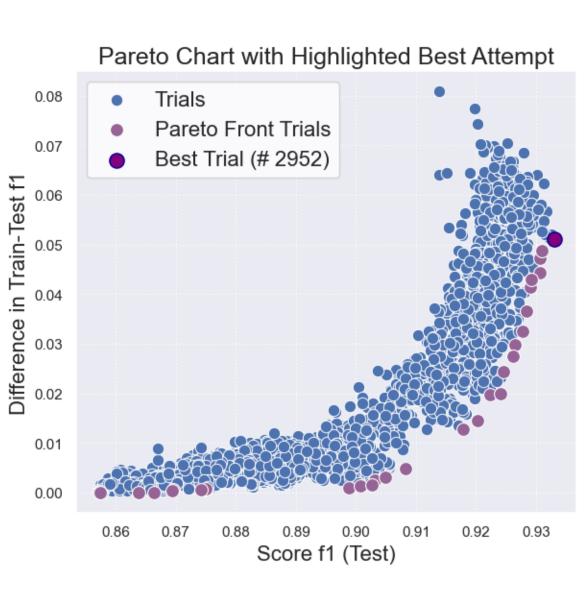
True label	40	000	- 150 0.8				, and a		
1-	12 0 Predicted	282 1 d label	n of positives						
	Confusion Ma	atrix (nomr.)	Fraction P.0						
label O	0.62	0.38	- 0.8 - 0.6 - 0.2						
True label	0.041	0.96	- 0.4 - 0.2 0.0	, e e e e e e e e e e e e e e e e e e e				train: 0.014 test: 0.0907	,
	0 Predicted	1 d label		0.0	0.2	0.4 Mean predicted prob	0.6 ability	0.8	1.0

70		Ac	ctive/Inactive	probability		
60	Ina	ctive zone			Active zone	type active (294) inactive (98) decoy (391)
50						
ercent						
20						
10						
0	0.0 0	.2).4	0.6	0.8	1.0
			Probabil	ity		

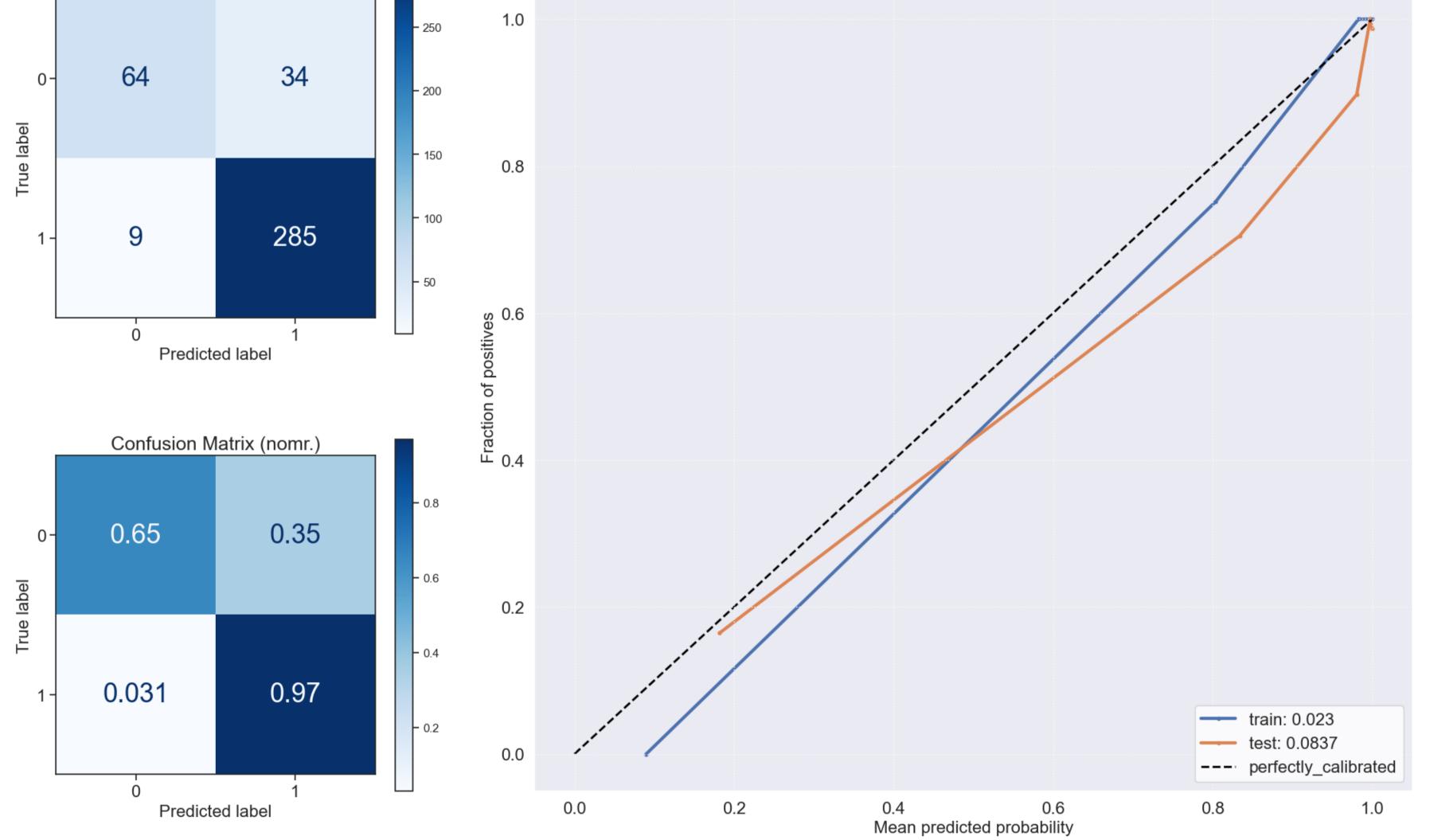
Metric	Train	Test
AUC	0.9993	0.9362
accuracy	0.9879	0.875
balanced accuracy	0.9758	0.7908
recall_(sens)	1.0	0.9592
specificity	0.9517	0.6224
precision	0.9841	0.884
f1_score	0.9878	0.8684

Protein: Q8NER1 - metric: f1





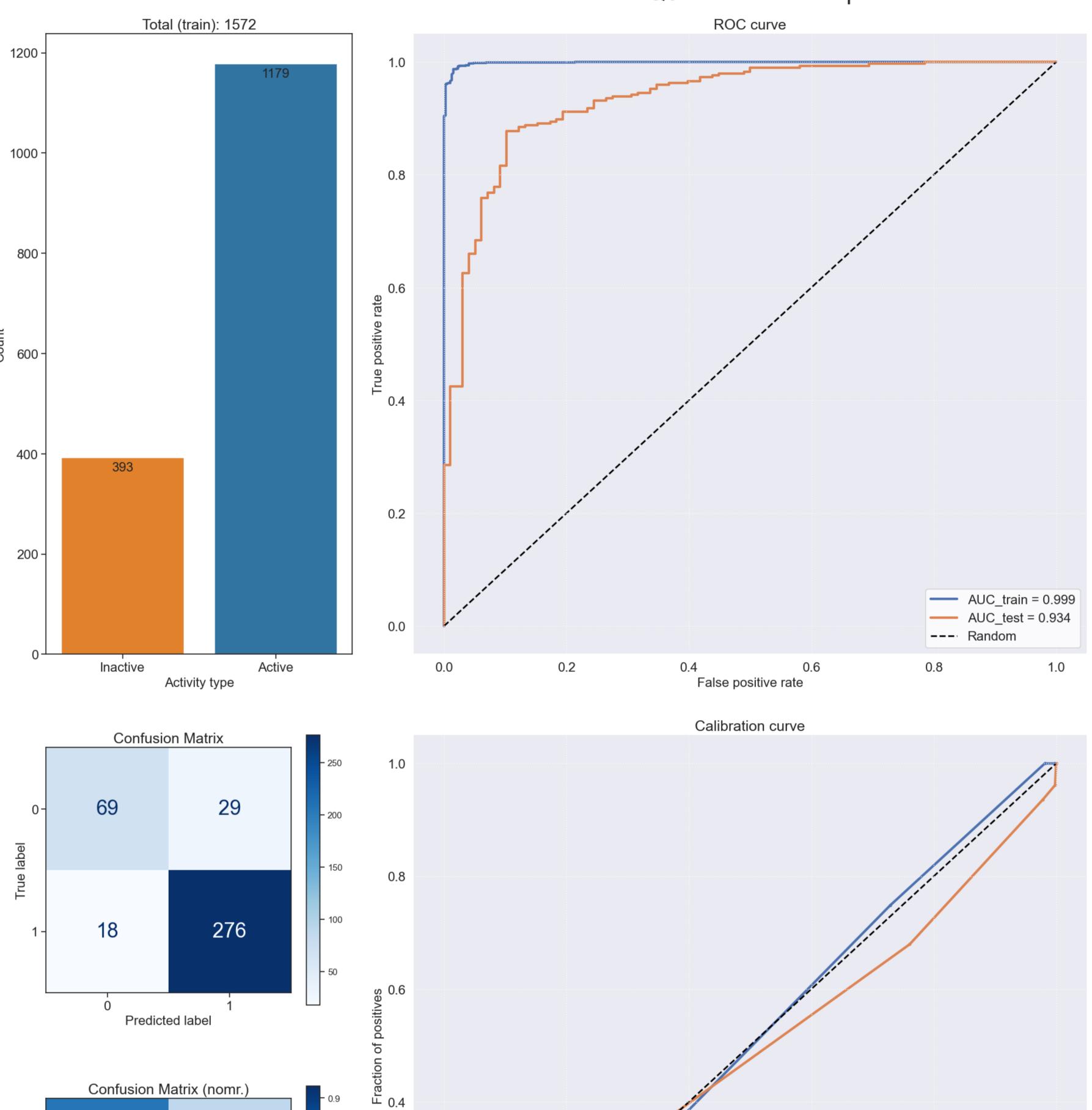
XGBoost Parameters	Value
alpha	0.0694
colsample_bytree	0.5036
gamma	0.2064
lambda	0.884
learning_rate	0.1478
max_depth	3.0
min_child_weight	2.0
n_estimators	765.0
reg_alpha	0.8449
reg_lambda	0.4548
scale_pos_weight	2.1901
subsample	0.7348

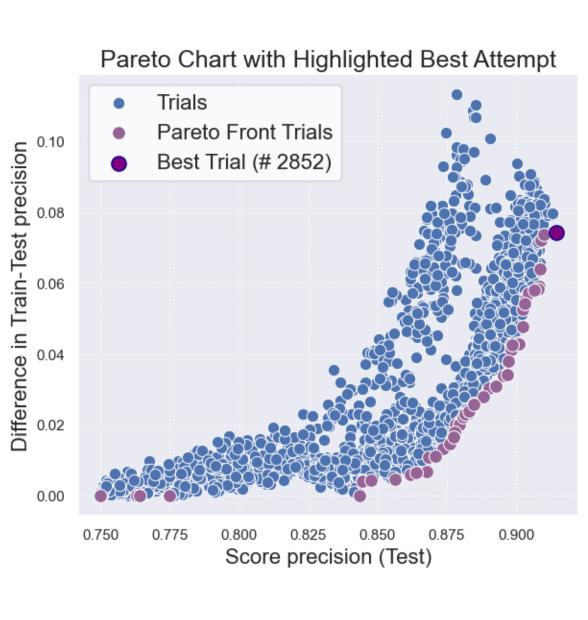


60 -			tive/Inactive				
	Ina	ctive zone			Active z	typ active inac	
40							
Percent							
20							
10							
0	0.0 0.	2 0	.4 Probabi	0.6	0.8	8	1.0

Metric	Train	Test
AUC	0.9977	0.9392
accuracy	0.9746	0.8903
balanced accuracy	0.9508	0.8112
recall_(sens)	0.9983	0.9694
specificity	0.9033	0.6531
precision	0.9687	0.8934
f1_score	0.9741	0.8845

Protein: Q8NER1 - metric: precision





XGBoost Parameters	Value
alpha	2.0487
colsample_bytree	0.5833
gamma	0.3573
lambda	3.9565
learning_rate	0.3353
max_depth	5.0
min_child_weight	3.0
n_estimators	677.0
reg_alpha	0.2368
reg_lambda	0.6842
scale_pos_weight	1.0591
subsample	0.8436

apel			- 0.6 - 0.5	2			
True label	0.061	0.94	- 0.5 - 0.4 - 0.3 - 0.2 - 0.1	, or or or or			
	0 Predicte	1 ed label	0.1	0.0	0.2	0.4 Mean predicted	0.6 probability
			Active/Inactive	ve probability			
60 -		Inactive zone			Active zone	type active (294) inactive (98) decoy (391)	
40							
Percent 08							

0.6

Probability

8.0

1.0

- 0.8

- 0.7

0.4

Predicted label

Confusion Matrix (nomr.)

0.3

0.2

0.7

0-

20 -

10 -

0.0

Metric	Train	Test
AUC	0.9989	0.9343
accuracy	0.986	0.8801
balanced accuracy	0.9779	0.8214
recall_(sens)	0.9941	0.9388
specificity	0.9618	0.7041
precision	0.9874	0.9049
f1_score	0.986	0.8776

AUC	0.9989	0.9343
accuracy	0.986	0.8801
balanced accuracy	0.9779	0.8214
recall_(sens)	0.9941	0.9388
specificity	0.9618	0.7041
precision	0.9874	0.9049
f1_score	0.986	0.8776

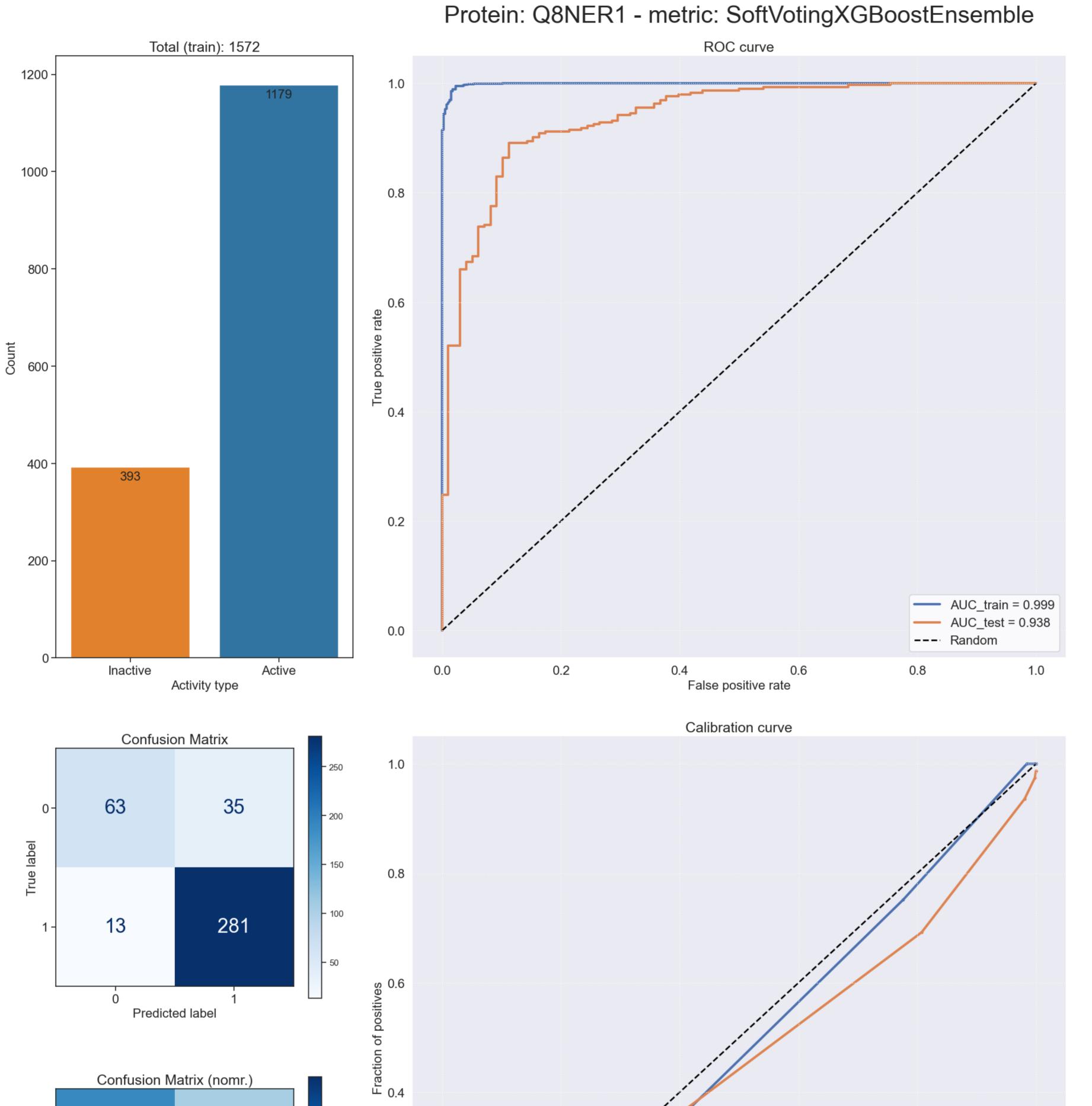
--- train: 0.0155

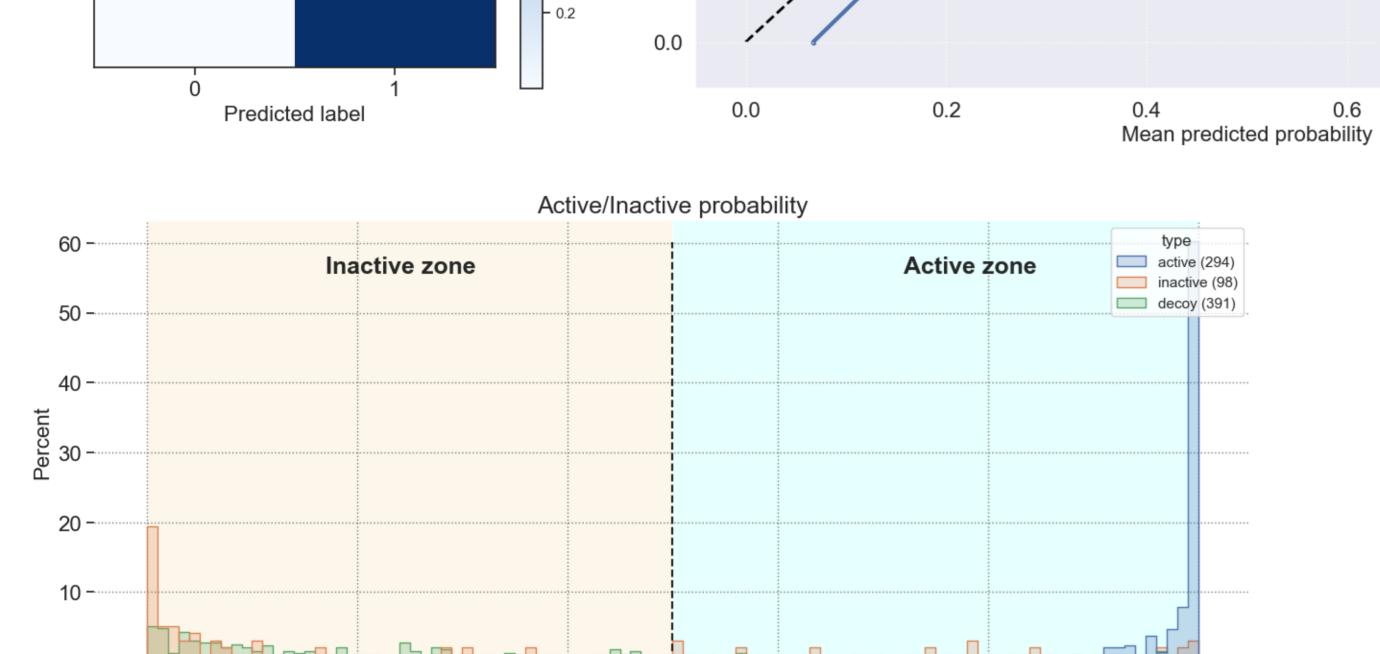
0.8

test: 0.0855

--- perfectly_calibrated

1.0





Probability

0.6

0.8

1.0

0.2

Confusion Matrix (nomr.)

0.36

0.96

0.2

0.6

0.4

0.4

0.64

0.044

0.0

0-

True label

Metric	Train	Test
AUC	0.999	0.938
accuracy	0.9841	0.8776
balanced accuracy	0.969	0.7993
recall_(sens)	0.9992	0.9558
specificity	0.9389	0.6429
precision	0.98	0.8892
f1_score	0.9839	0.872

train: 0.0165

test: 0.0844

8.0

perfectly_calibrated

1.0

0.8

0.6

0.4

0.2

0.0

0.0

0.2

0.6