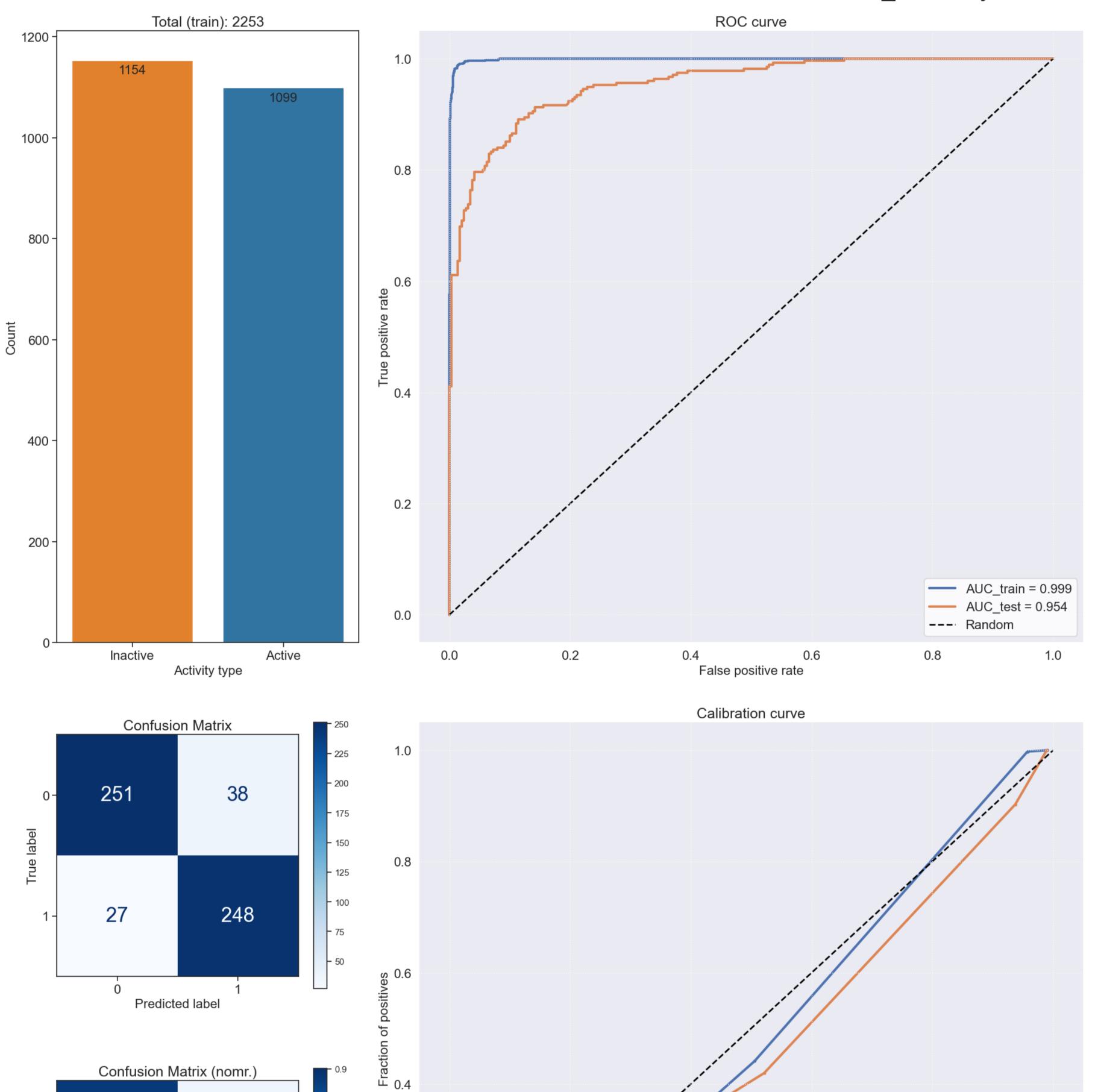
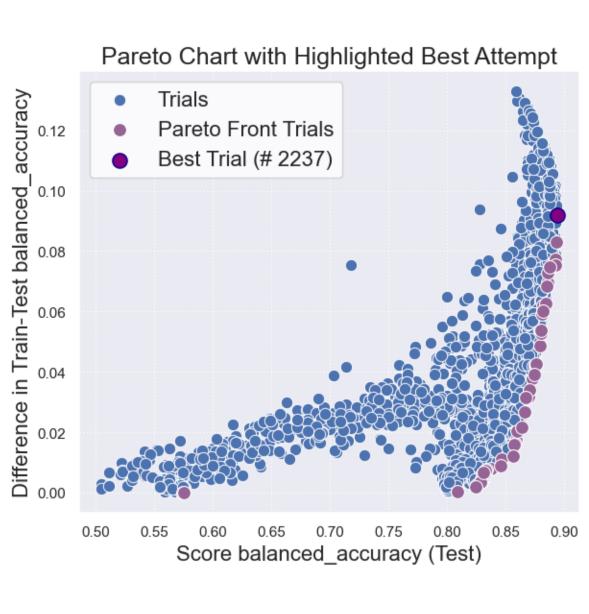
Protein: P50406 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	2.2117
colsample_bytree	0.5027
gamma	0.5821
lambda	2.573
learning_rate	0.0399
max_depth	7.0
min_child_weight	1.0
n_estimators	651.0
reg_alpha	0.0408
reg_lambda	0.877
scale_pos_weight	1.4495
subsample	0.8213

	Predicted label	0.1	0.0	0.2	0.4 Mean predicted p	0.6 robability
		Active/In	active probability			
40	Inac	tive zone		Active zone	type active (275) inactive (289) decoy (563)	
Percent - 05						
10 						
	0.0 0.2		0.6 Probability	0.8	1.0	

0.2

0.0

- 0.8

0.4

0.3

0.2

Confusion Matrix (nomr.)

0.13

0.9

0.87

0.098

0-

Metric	Train	Test
AUC	0.9988	0.9543
accuracy	0.984	0.8848
balanced accuracy	0.9843	0.8852
recall_(sens)	0.9945	0.9018
specificity	0.974	0.8685
precision	0.9733	0.8671
f1_score	0.984	0.8848

--- train: 0.0188

0.8

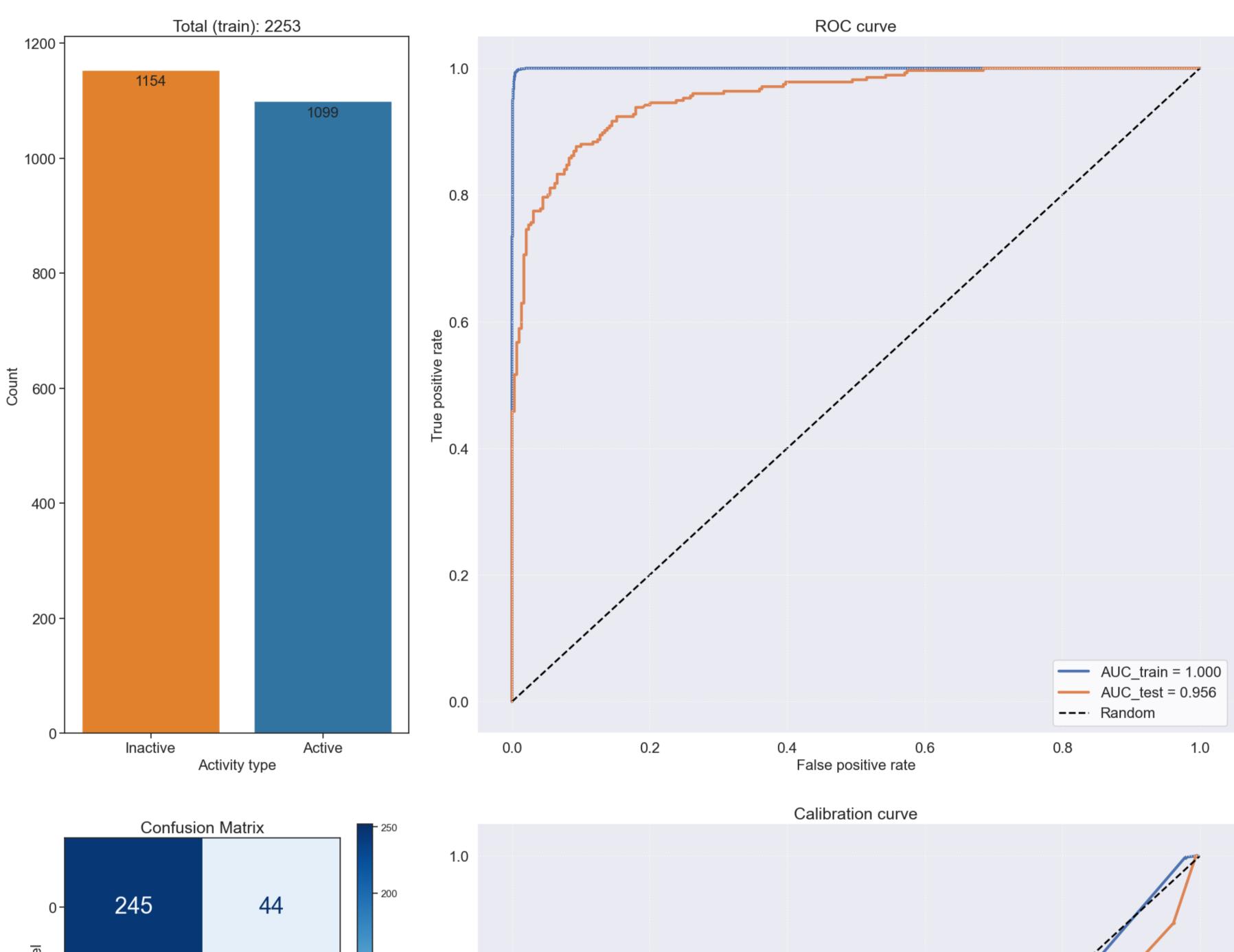
test: 0.0874

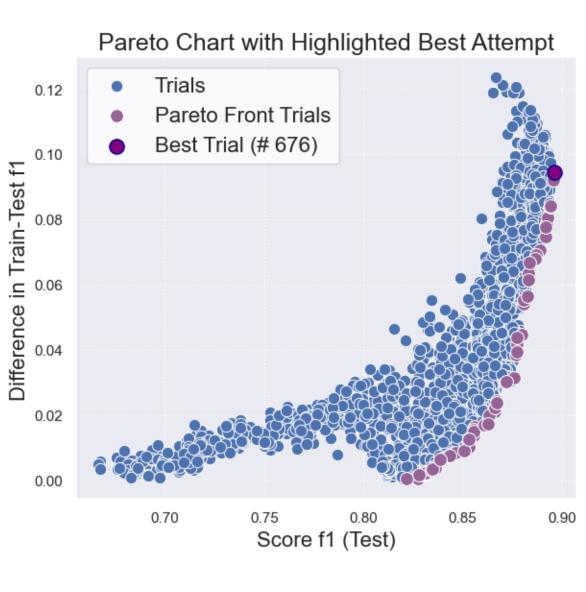
--- perfectly_calibrated

1.0

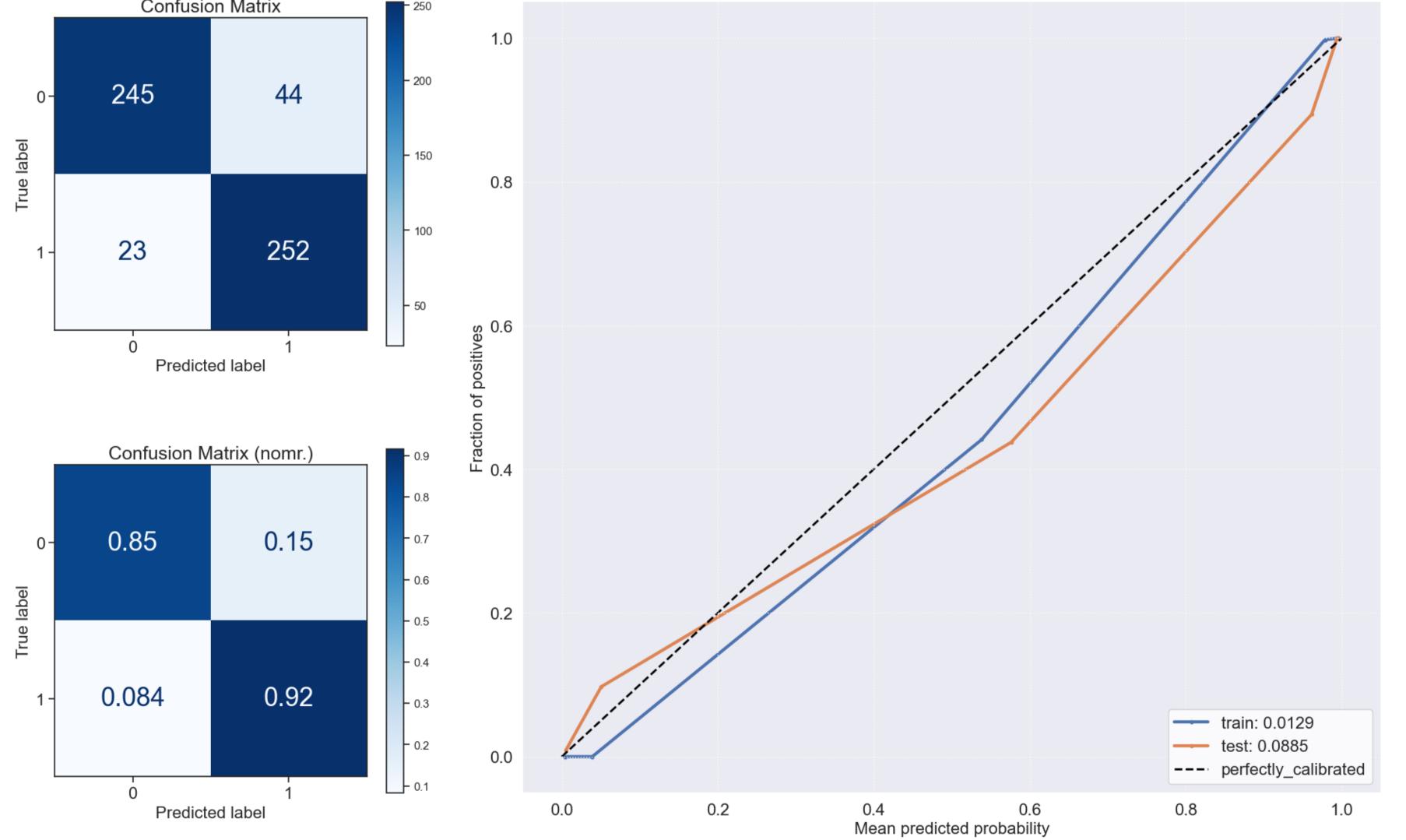
		Activ	/e/Inactive pro	bability			
40	Inac	ctive zone			Active zone	type active (275) inactive (28	39)
						decoy (563)	
30 - 20 - 20 - 20 - 20 - 20 - 20 - 20 -						Π	
0							
0.0	0.2	0.4	Probability	0.6	0.8	1.0	

Protein: P50406 - metric: f1





XGBoost Parameters	Value
alpha	4.3648
colsample_bytree	0.5091
gamma	0.2546
lambda	3.2404
learning_rate	0.0275
max_depth	10.0
min_child_weight	1.0
n_estimators	682.0
reg_alpha	0.5273
reg_lambda	0.517
scale_pos_weight	2.5829
subsample	0.8726

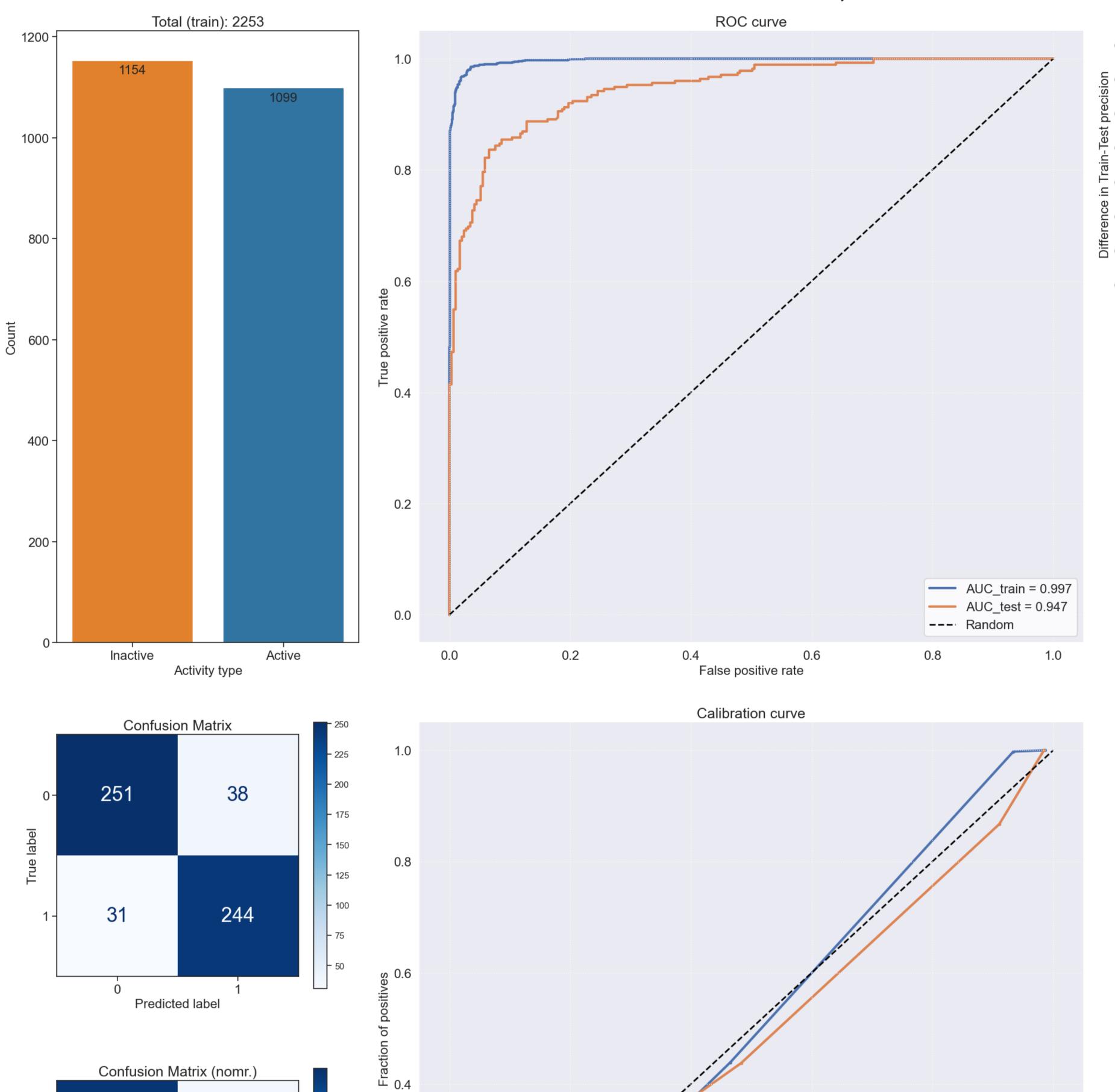


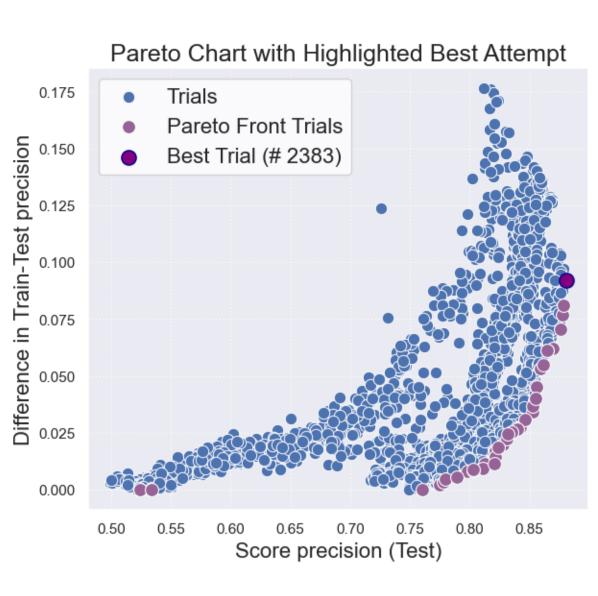
40	I I	ictive zone	1		Active z	one a	ype ctive (275) active (289)
						de	active (289) ecoy (563)
Percent - 05							
10							
0	.0 0	.2	0.4 Probabi	0.6 lity	0.8	8	1.0

Active/Inactive probability

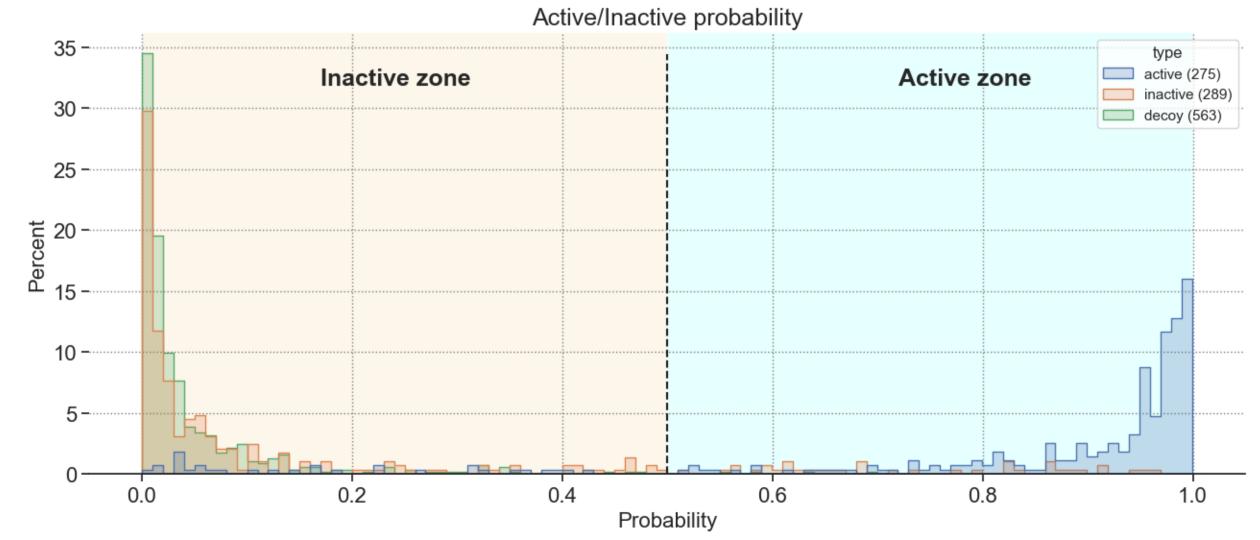
Metric	Train	Test
AUC	0.9996	0.9558
accuracy	0.9902	0.8812
balanced accuracy	0.9904	0.8821
recall_(sens)	0.9991	0.9164
specificity	0.9818	0.8478
precision	0.9812	0.8514
f1_score	0.9902	0.8812

Protein: P50406 - metric: precision





XGBoost Parameters	Value
alpha	4.8015
colsample_bytree	0.9801
gamma	0.2199
lambda	0.3479
learning_rate	0.0305
max_depth	6.0
min_child_weight	3.0
n_estimators	936.0
reg_alpha	0.1438
reg_lambda	0.5345
scale_pos_weight	1.0285
subsample	0.7246



Confusion Matrix (nomr.)

Predicted label

0.13

0.89

0.6

0.4

- 0.3

0.2

0.2

0.0

0.0

0.2

0.87

0.11

0-

True label

Metric	Train	Test
AUC	0.9967	0.9467
accuracy	0.9734	0.8777
balanced accuracy	0.9734	0.8779
recall_(sens)	0.9763	0.8873
specificity	0.9705	0.8685
precision	0.9693	0.8652
f1_score	0.9734	0.8777

- train: 0.0263

0.8

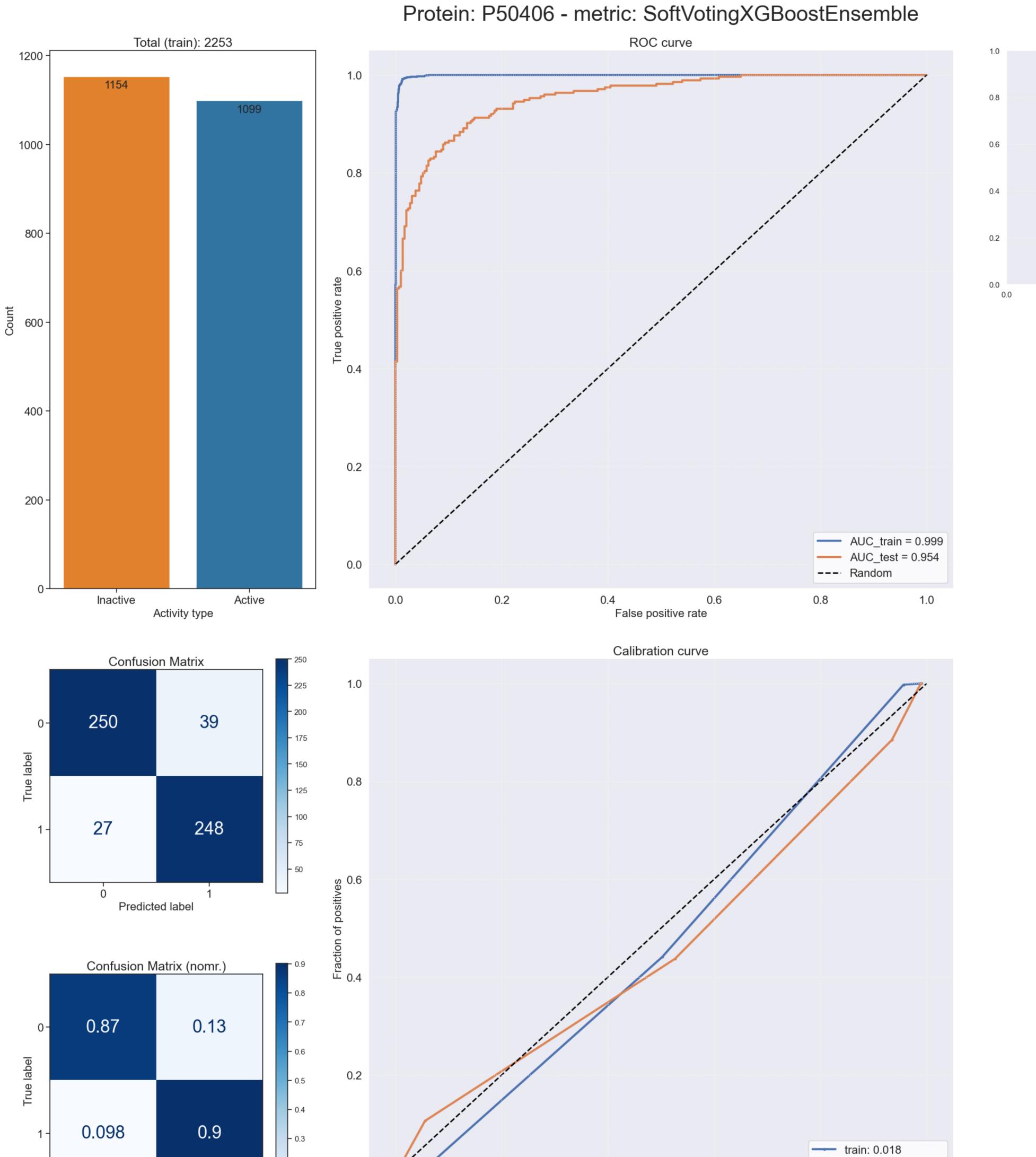
test: 0.0926

--- perfectly_calibrated

1.0

Active/iii	lactive probability				
Inactive zone	Active zone	type active (275)	Metric	Train	Test
		inactive (289) decoy (563)	AUC	0.9967	0.94
			accuracy	0.9734	0.87
1			balanced accuracy	0.9734	0.87
			recall_(sens)	0.9763	0.88
			specificity	0.9705	0.86
			precision	0.9693	0.86
			f1_score	0.9734	0.87
0.2 0.4	0.6 0.8	1.0			

0.4 0.6 Mean predicted probability



0.2

0.4 0.6 Mean predicted probability

Metric	Train	Test
AUC	0.999	0.9539
accuracy	0.9849	0.883
balanced accuracy	0.9852	0.8834
recall_(sens)	0.9955	0.9018
specificity	0.9749	0.8651
precision	0.9742	0.8641
f1_score	0.9849	0.883

test: 0.0878

8.0

--- perfectly_calibrated

1.0

0.2

0.6

			Active	/Inactive proba	ability		
35		Inactive zone				Active zone	active (275)
30							
Percent 25							П
20 15							
10							
5-·····							
Ü	0.0	0.2	0.4	Probability	0.6	0.8	1.0

0.2

Predicted label

0.0

0.0