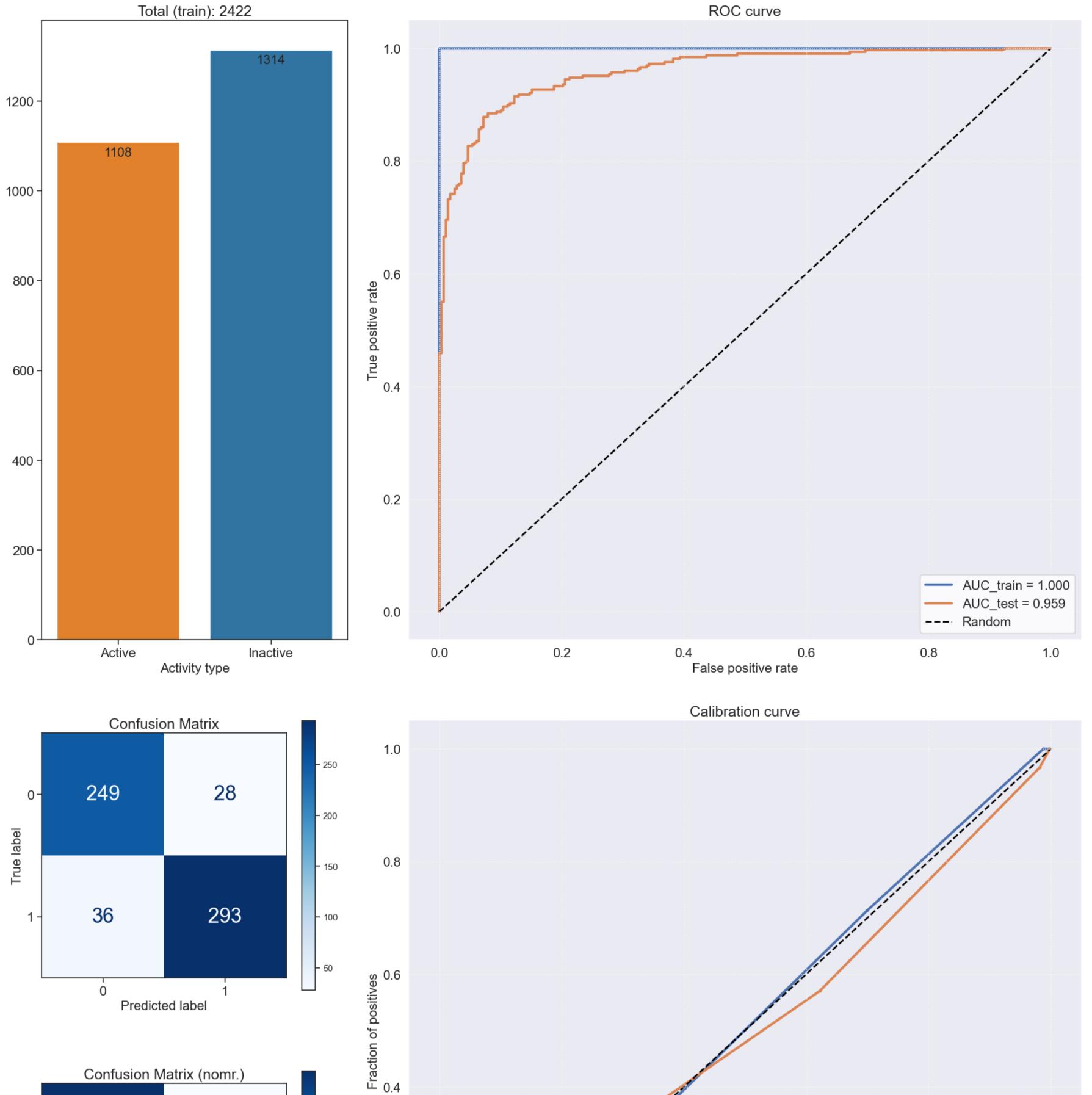
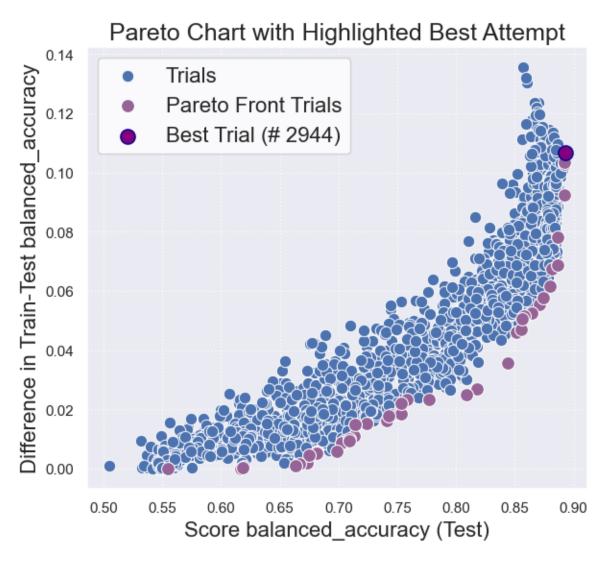
Protein: Q16539 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	1.7039
colsample_bytree	0.6115
gamma	0.0882
lambda	0.3286
learning_rate	0.0653
max_depth	9.0
min_child_weight	1.0
n_estimators	786.0
reg_alpha	0.3519
reg_lambda	0.2733
scale_pos_weight	1.0653
subsample	0.5434

Active/Inactive probability							
50 		Inactive zone			Active z	inac	ve (329) ctive (277) oy (604)
20							
Percent							
10							
0-	0.0	0.2	0.4	0.6	0.	8	1.0
			Proi	pability			

0.2

0.0

0.0

0.2

0.4

0.3

0.2

Confusion Matrix (nomr.)

Predicted label

0.9

0.11

0-

True label

0.1

0.89

Metric	Train	Test
AUC	1.0	0.9593
accuracy	0.9996	0.8944
balanced accuracy	0.9995	0.8947
recall_(sens)	1.0	0.8906
specificity	0.9991	0.8989
precision	0.9992	0.9128
f1_score	0.9996	0.8945

Metric	Train	Test
AUC	1.0	0.9593
accuracy	0.9996	0.8944
balanced accuracy	0.9995	0.8947
recall_(sens)	1.0	0.8906
specificity	0.9991	0.8989
precision	0.9992	0.9128
f1_score	0.9996	0.8945

--- train: 0.0042

0.8

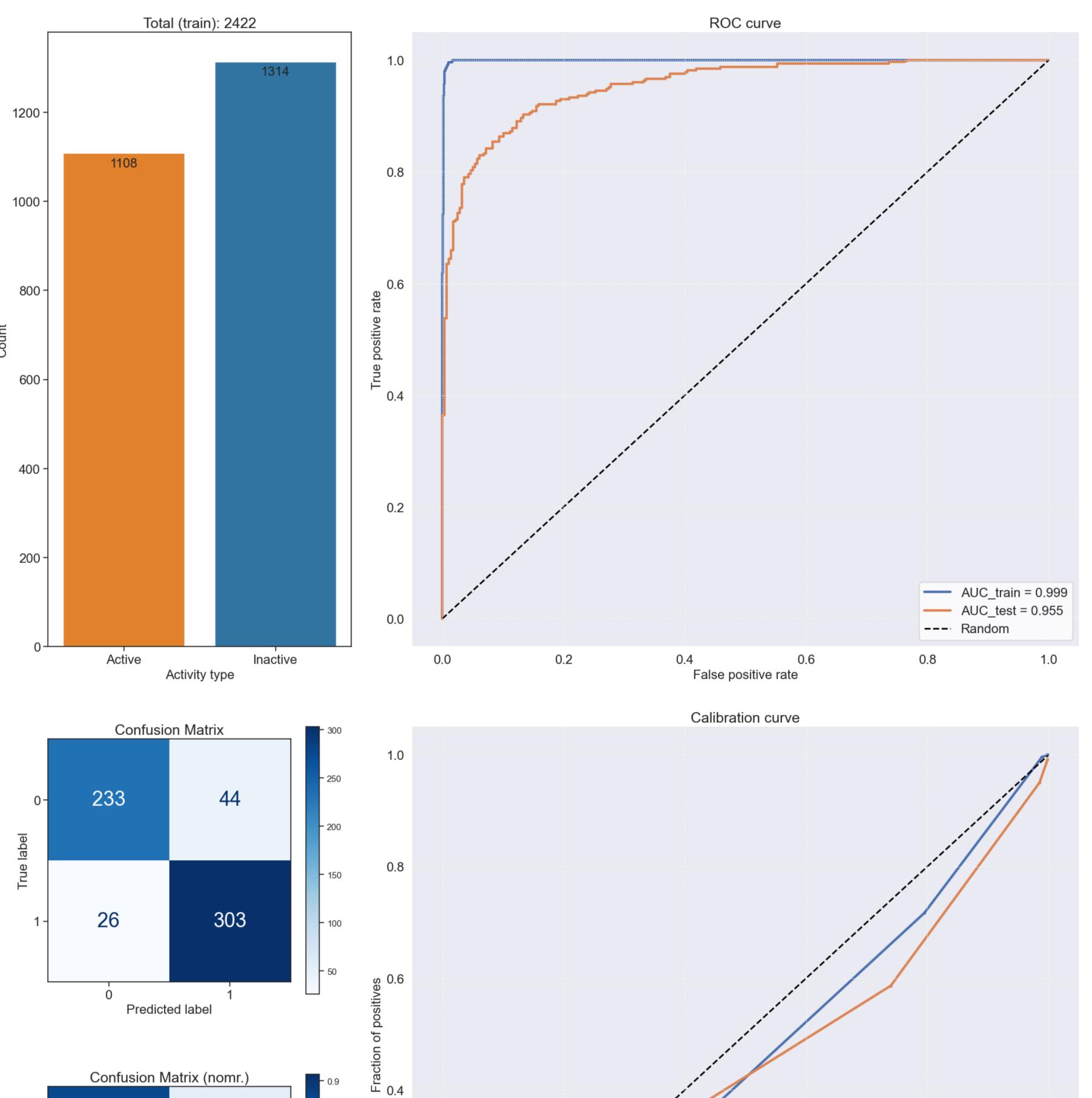
0.4 0.6 Mean predicted probability

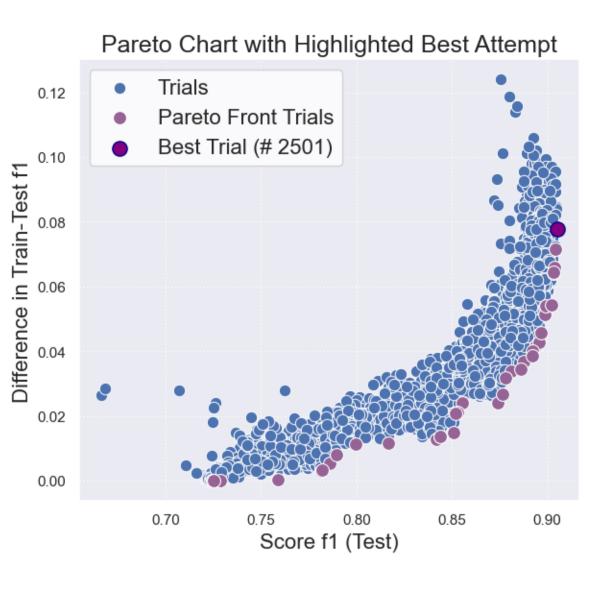
test: 0.0796

--- perfectly_calibrated

1.0

Protein: Q16539 - metric: f1





XGBoost Parameters	Value
alpha	4.92
colsample_bytree	0.7754
gamma	0.01
lambda	2.2124
learning_rate	0.0686
max_depth	10.0
min_child_weight	10.0
n_estimators	863.0
reg_alpha	0.1049
reg_lambda	0.634
scale_pos_weight	3.4759
subsample	0.8996

		Δ	ctive/Inactiv	e probabilit	У		
		Inactive zone			Active	zone	type active (329) inactive (277) decoy (604)
40							
 96							
10							
0	0.0	0.2	0.4	0.6	0	.8	1.0
			Proba	ability			

Confusion Matrix (nomr.)

Predicted label

0.16

0.92

- 0.6

- 0.5

0.4

0.3

0.2

0.2

0.0

0.0

0.2

0.84

0.079

0-

True label

Metric	Train	Test
AUC	0.9992	0.9546
accuracy	0.9818	0.8845
balanced accuracy	0.9801	0.8811
recall_(sens)	1.0	0.921
specificity	0.9603	0.8412
precision	0.9676	0.8732
f1_score	0.9818	0.8841

--- train: 0.0169

0.8

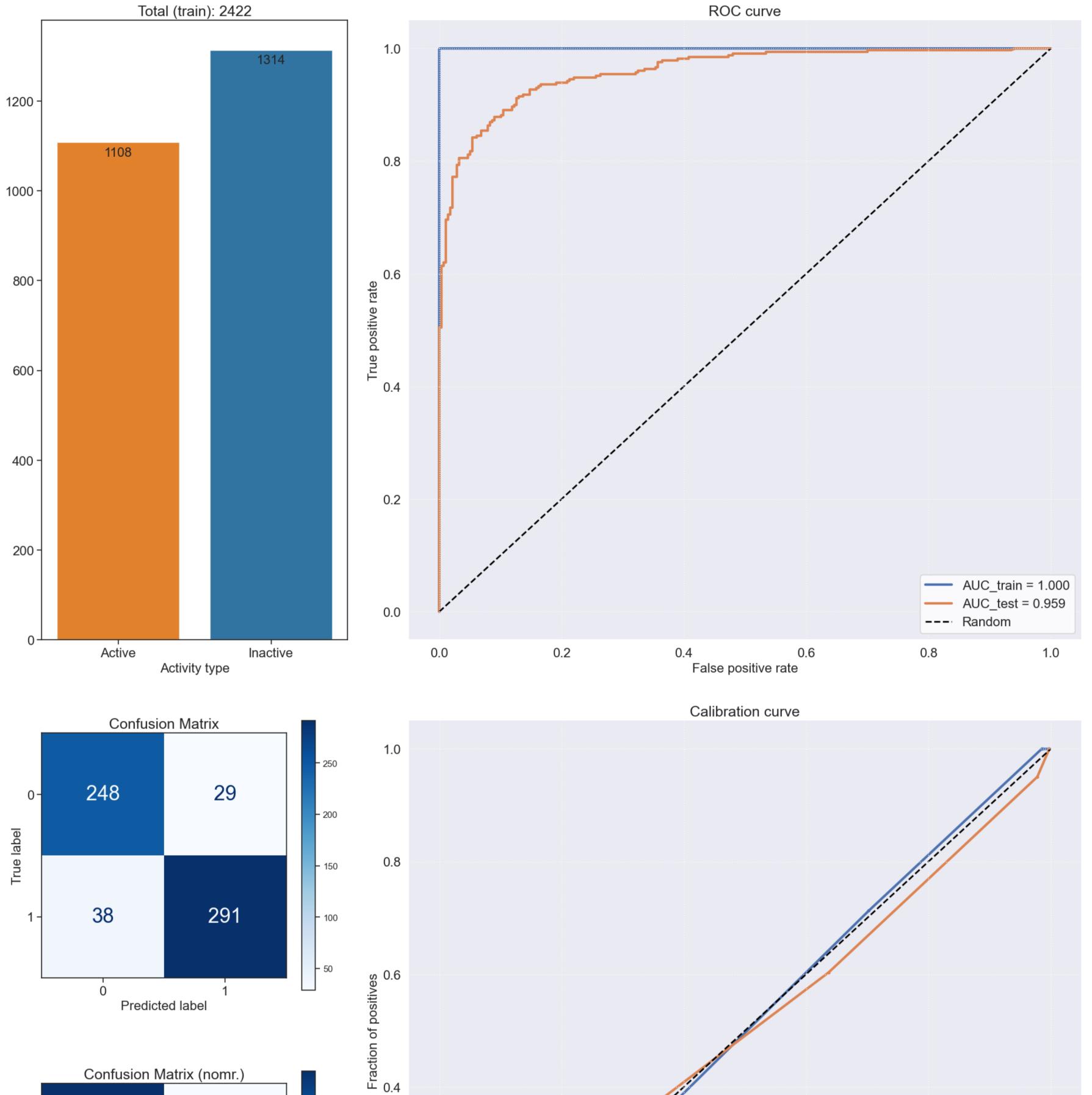
0.4 0.6 Mean predicted probability

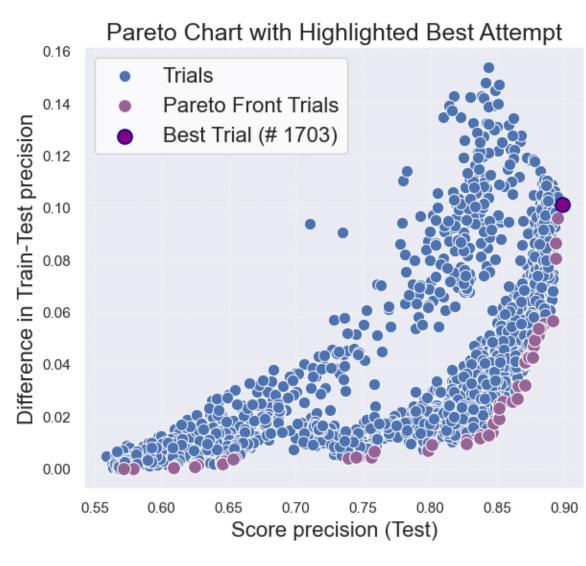
test: 0.0895

--- perfectly_calibrated

1.0

Protein: Q16539 - metric: precision





XGBoost Parameters	Value
alpha	3.2136
colsample_bytree	0.6343
gamma	0.273
lambda	2.392
learning_rate	0.0336
max_depth	10.0
min_child_weight	1.0
n_estimators	987.0
reg_alpha	0.1503
reg_lambda	0.2126
scale_pos_weight	1.1392
subsample	0.6139

			Active/Inactive	ve probability		
40		Inactive zone			Active zone	
30						decoy (604)
Percent 00						
40						
0-						
0	0.0	0.2	0.4 Proba	0.6 ability	0.8	1.0

0.2

0.0

0.0

0.2

0.5

0.4

0.3

0.2

Confusion Matrix (nomr.)

Predicted label

0.1

0.88

0.9

0.12

0-

True label

Metric	Train	Test
AUC	1.0	0.9594
accuracy	0.9992	0.8894
balanced accuracy	0.9992	0.8899
recall_(sens)	0.9992	0.8845
specificity	0.9991	0.8953
precision	0.9992	0.9094
f1_score	0.9992	0.8896

zone	Active	type active (Metric	Train	Test
		inactive decoy	•	AUC	1.0	0.9594
				accuracy	0.9992	0.8894
				balanced accuracy	0.9992	0.8899
				recall_(sens)	0.9992	0.8845
				specificity	0.9991	0.8953
				precision	0.9992	0.9094
				f1_score	0.9992	0.8896
0.4	0.6	0.8 1	0			

0.4 0.6 Mean predicted probability

--- train: 0.0055

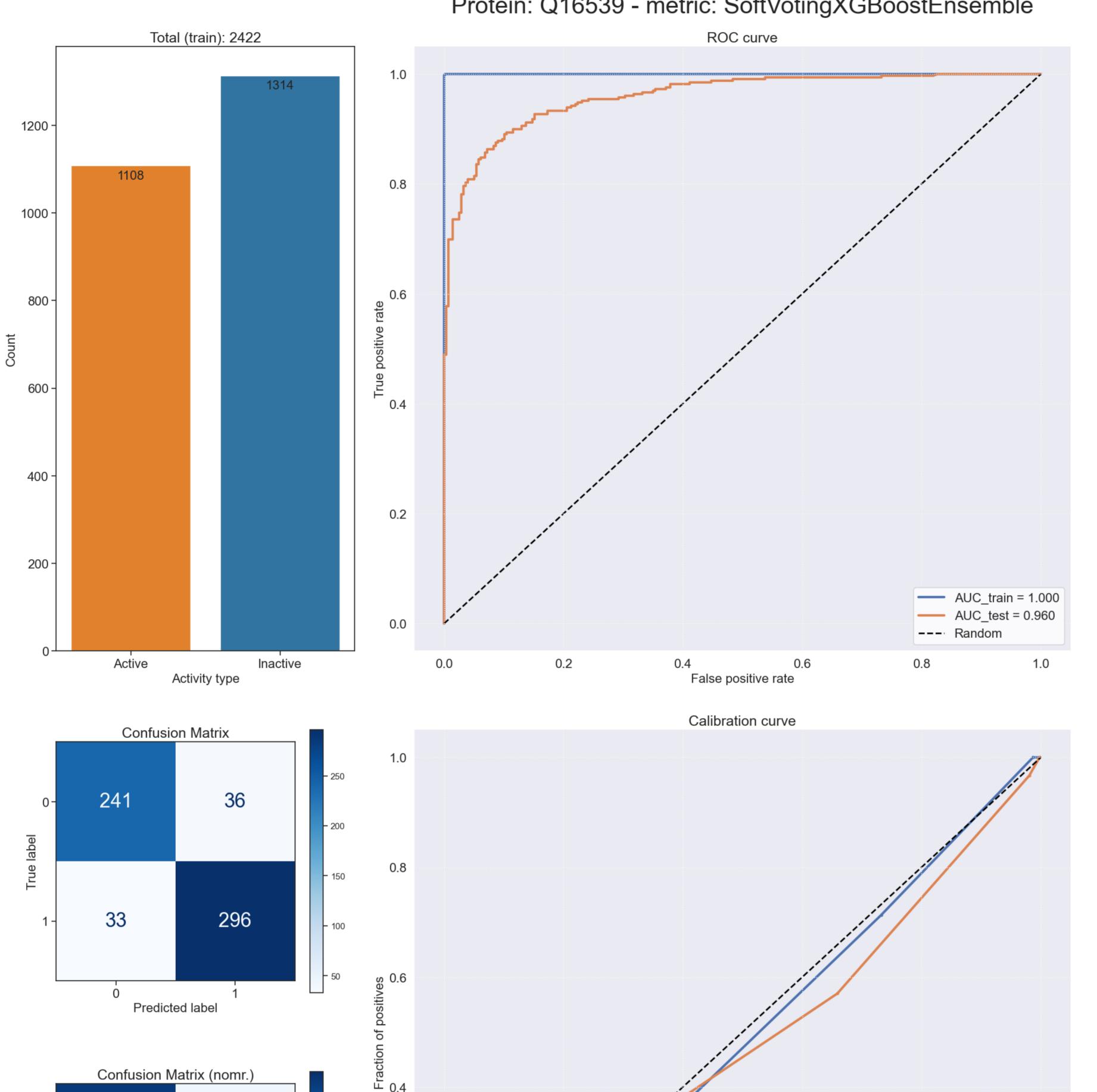
0.8

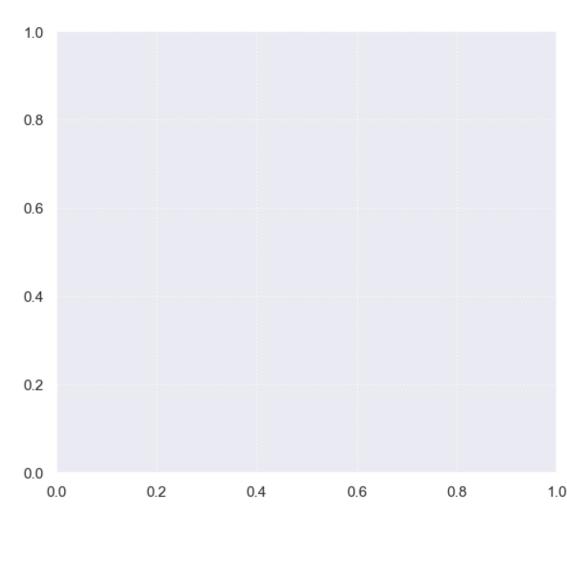
test: 0.0804

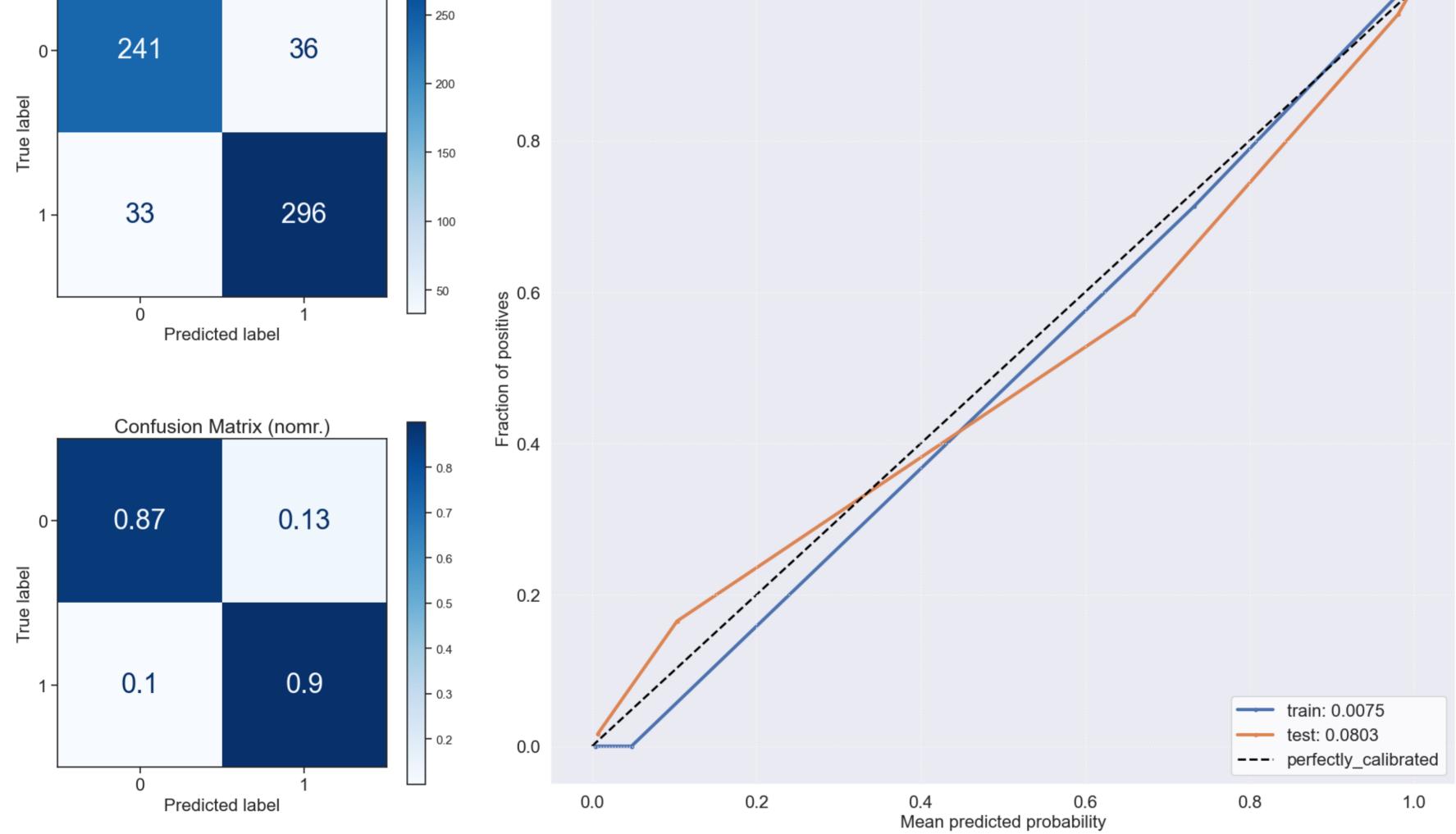
--- perfectly_calibrated

1.0

Protein: Q16539 - metric: SoftVotingXGBoostEnsemble







F0 -		Ac	tive/Inactiv	e probabi	lity		
40	Inact	tive zone			Active		type active (329) inactive (277) decoy (604)
± 30							
Percent 00							
10							
0	0.0 0.2	0	.4	0.	.6 0	0.8	1.0
			Proba				

Metric	Train	Test
AUC	1.0	0.9598
accuracy	0.9959	0.8861
balanced accuracy	0.9955	0.8849
recall_(sens)	1.0	0.8997
specificity	0.991	0.87
precision	0.9924	0.8916
f1_score	0.9959	0.8861