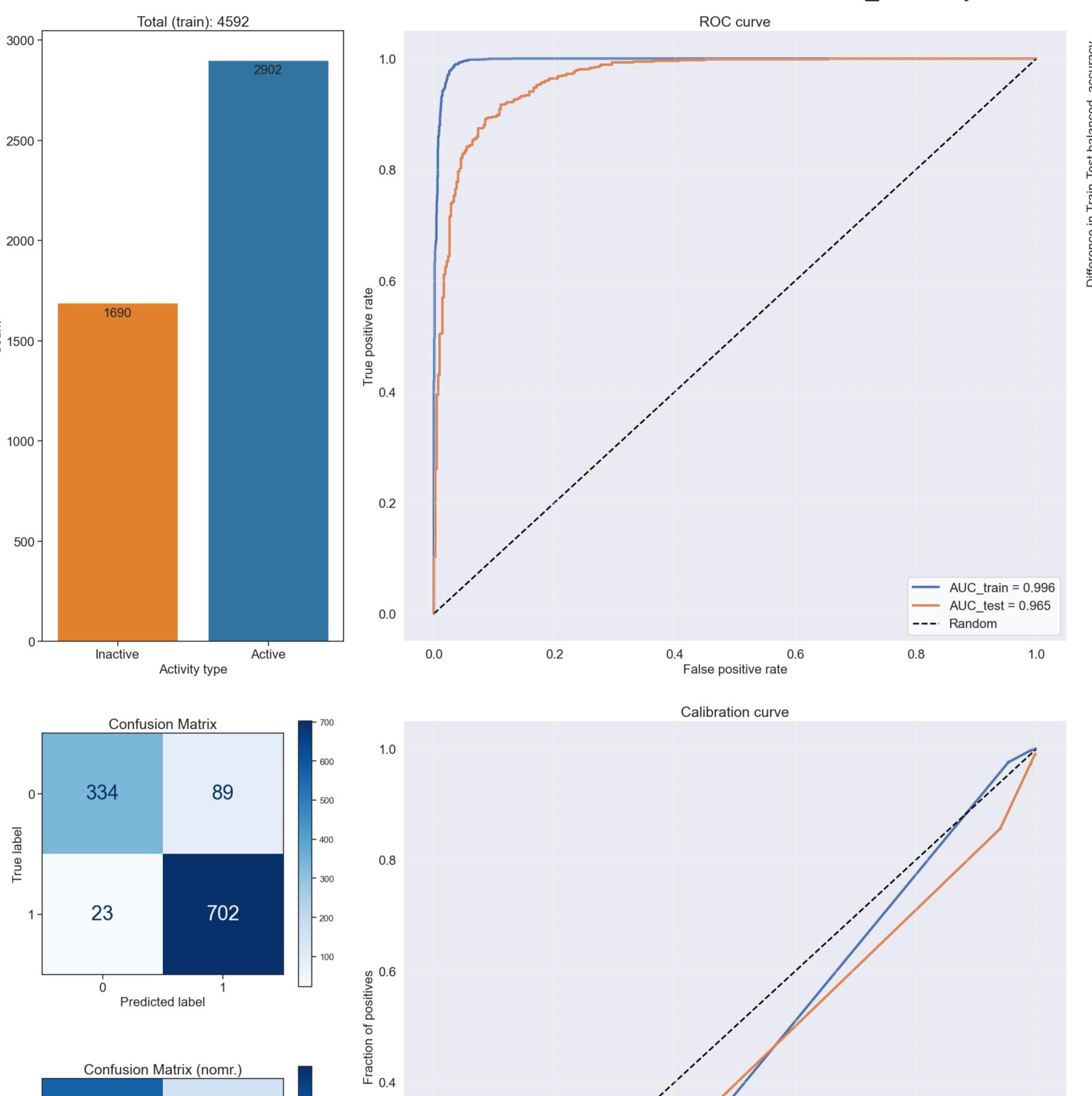
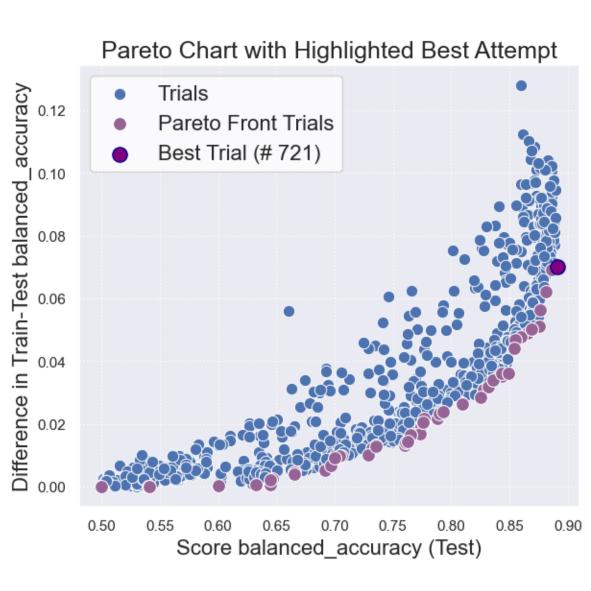
Protein: P56817 - metric: balanced_accuracy



0.2



XGBoost Parameters	Value
alpha	2.4795
colsample_bytree	0.5157
gamma	0.4708
lambda	2.2155
learning_rate	0.0721
max_depth	9.0
min_child_weight	3.0
n_estimators	569.0
reg_alpha	0.4314
reg_lambda	0.8586
scale_pos_weight	2.7138
subsample	0.5553

		Ac	tive/Inactive probab	ility		
40		ctive zone		Active	zone active (inactive decoy (e (423)
00						
Percent 00 00 00 00 00 00 00 00 00 00 00 00 00						
10						
0	0.0 0.	2 0	.4 0 Probability	.6 0	.8 1	.0

0.2

0.0

0.0

0.4

0.2

Confusion Matrix (nomr.)

Predicted label

0.21

0.97

0.79

0.032

0-

True label

Metric	Train	Test
AUC	0.9961	0.9649
accuracy	0.9713	0.9024
balanced accuracy	0.9614	0.8789
recall_(sens)	0.9986	0.9683
specificity	0.9243	0.7896
precision	0.9577	0.8875
f1_score	0.971	0.9004

--- train: 0.0252

0.8

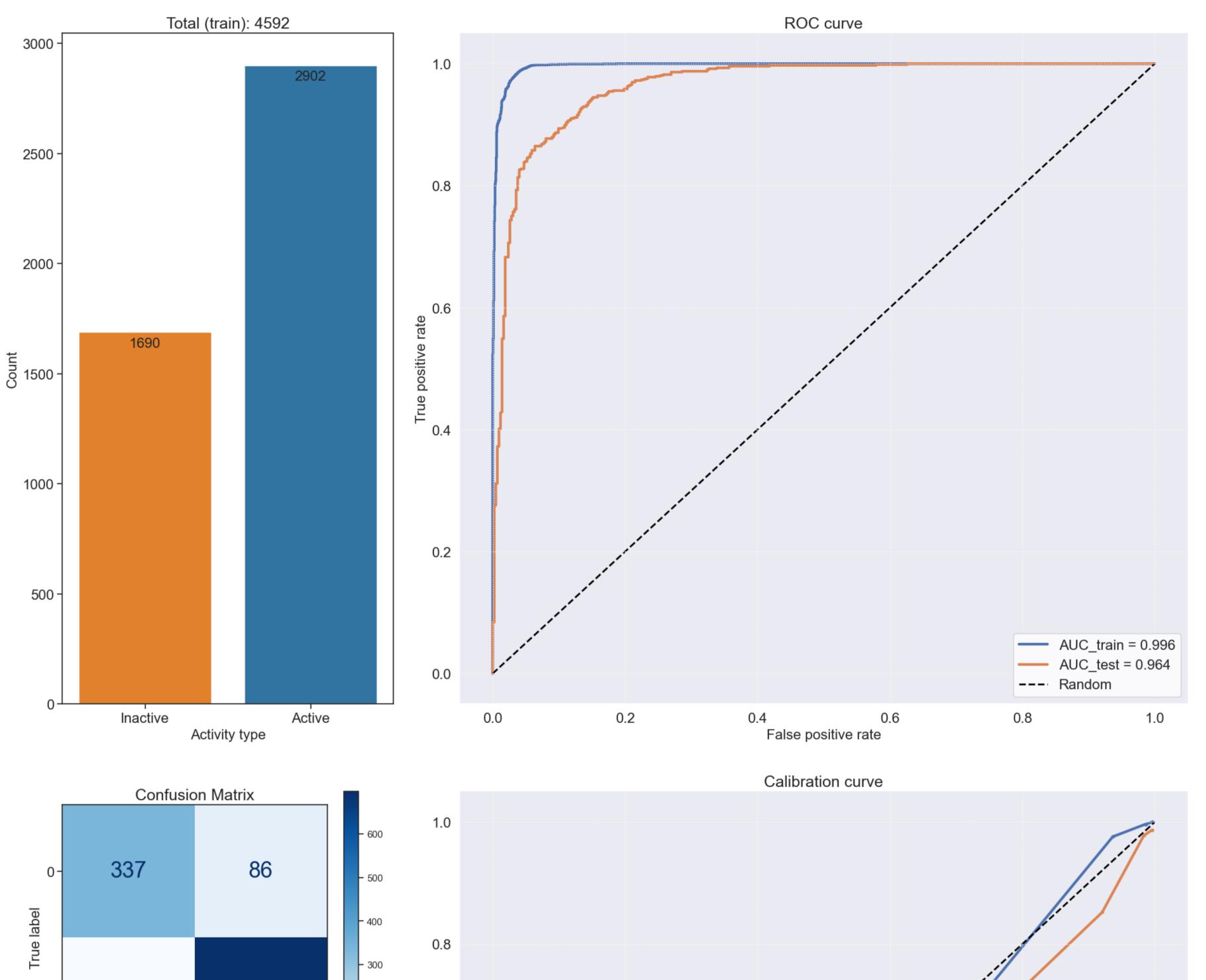
0.4 0.6 Mean predicted probability

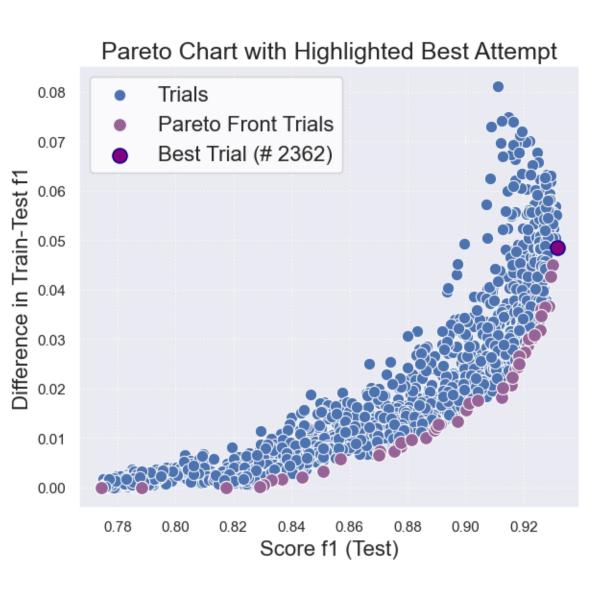
- test: 0.0718

--- perfectly_calibrated

1.0

Protein: P56817 - metric: f1





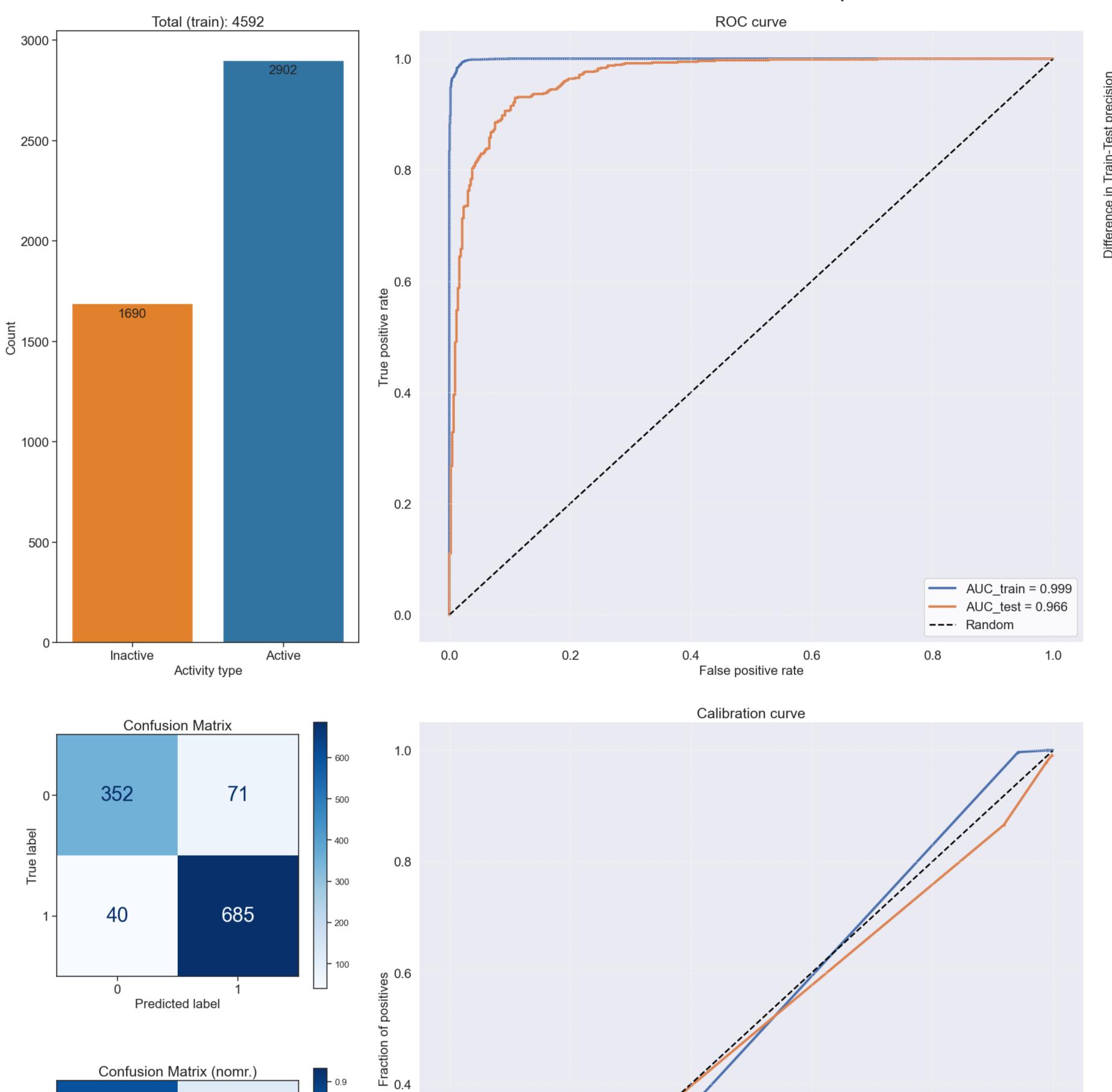
XGBoost Parameters	Value
alpha	1.4841
colsample_bytree	0.6248
gamma	0.0374
lambda	1.4728
learning_rate	0.0294
max_depth	8.0
min_child_weight	1.0
n_estimators	714.0
reg_alpha	0.8102
reg_lambda	0.3176
scale_pos_weight	1.9953
subsample	0.6906

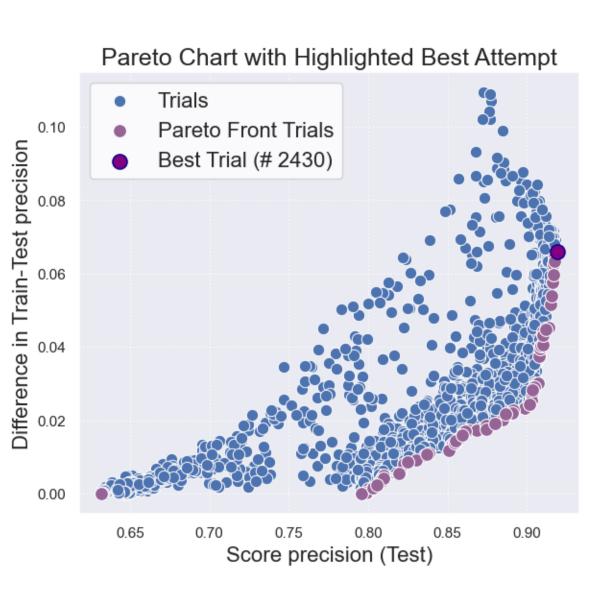
0-	337	86	- 500							
True label			- 400							
True			- 300	0.8					.://	
1-	28	697	- 200							
			- 100	0.0						
ı	0 Predicte	1 ed label		n of positives			, para range			
	Confusion M	latrix (nomr.)	- I	Fraction 4.0						
label O	0.8	0.2	- 0.8 - 0.6	0.2		, or				
True label	•	0.96	- 0.4 - 0.2	0.0					train: 0.024 test: 0.0706	;
	0 Predicte	1 ed label			0.0	0.2	0.4 Mean predicted pro	0.6 bability	0.8	1.0

40 -			ve probability		
0.5	Inactive zone			Active zone	type active (725)
30 -					decoy (1143)
20					
10					
5-					
0		0.4			4.0
0.0	U.Z	0.4 Prob	0.6 pability	0.δ	1.0

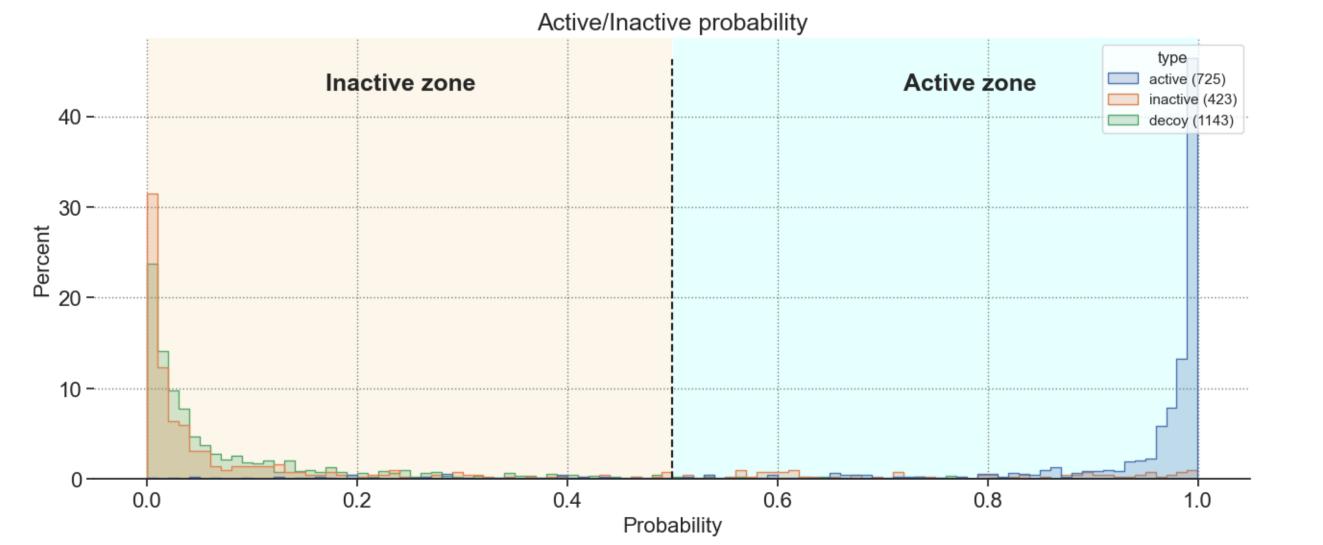
Metric	Train	Test
AUC	0.9965	0.9643
accuracy	0.9728	0.9007
balanced accuracy	0.9638	0.879
recall_(sens)	0.9979	0.9614
specificity	0.9296	0.7967
precision	0.9605	0.8902
f1_score	0.9726	0.899

Protein: P56817 - metric: precision





XGBoost Parameters	Value
alpha	1.0545
colsample_bytree	0.5502
gamma	0.1833
lambda	1.8003
learning_rate	0.0579
max_depth	10.0
min_child_weight	1.0
n_estimators	632.0
reg_alpha	0.5484
reg_lambda	0.3758
scale_pos_weight	1.118
subsample	0.8801



- 0.8

- 0.6

- 0.5

0.4

0.3

0.2

0.1

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.17

0.94

0.83

0.055

0-

True label

Metric	Train	Test
AUC	0.9992	0.9658
accuracy	0.9869	0.9033
balanced accuracy	0.9831	0.8885
recall_(sens)	0.9976	0.9448
specificity	0.9686	0.8322
precision	0.982	0.9061
f1_score	0.9869	0.9025

-- train: 0.0128

0.8

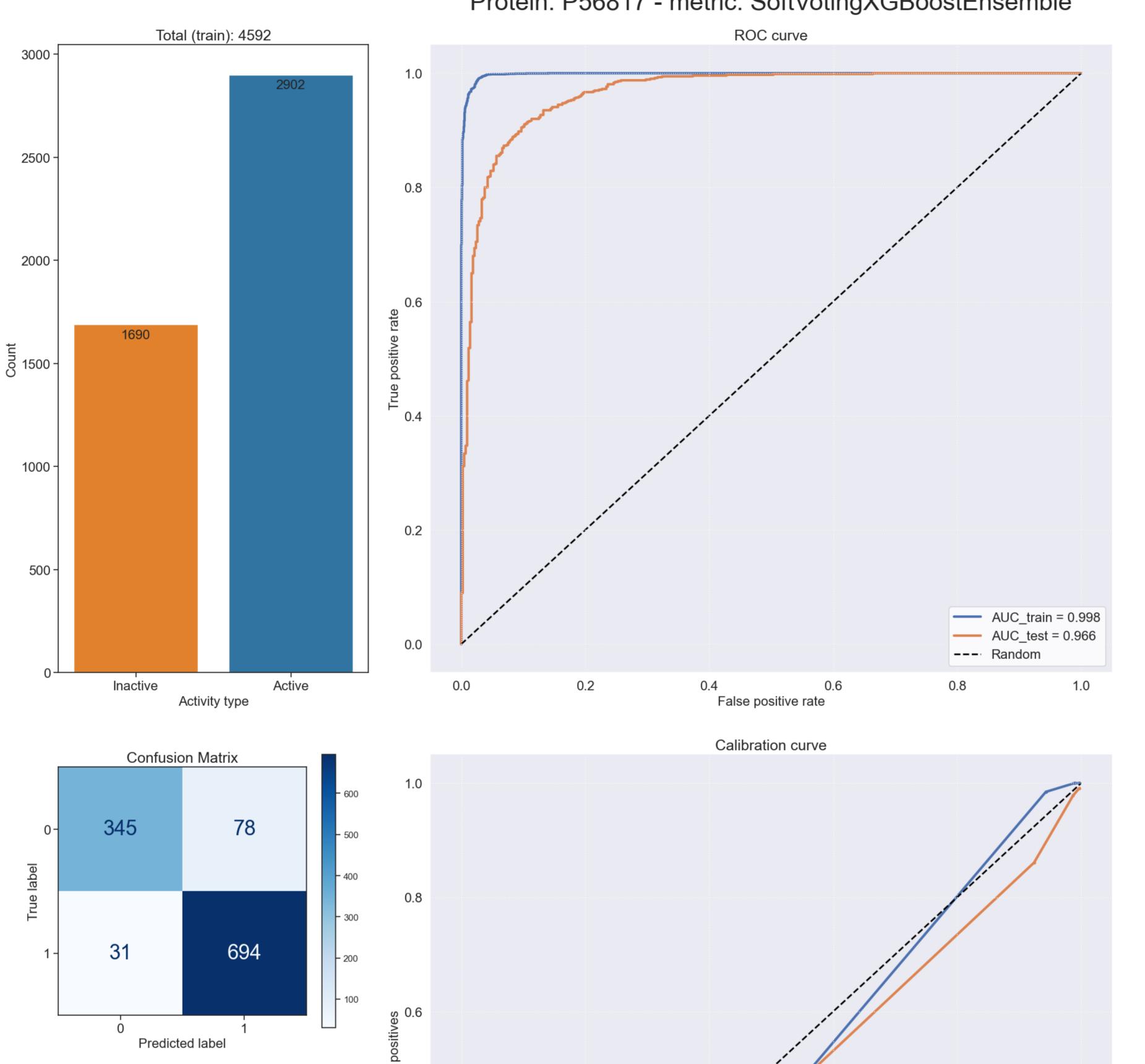
0.4 0.6 Mean predicted probability

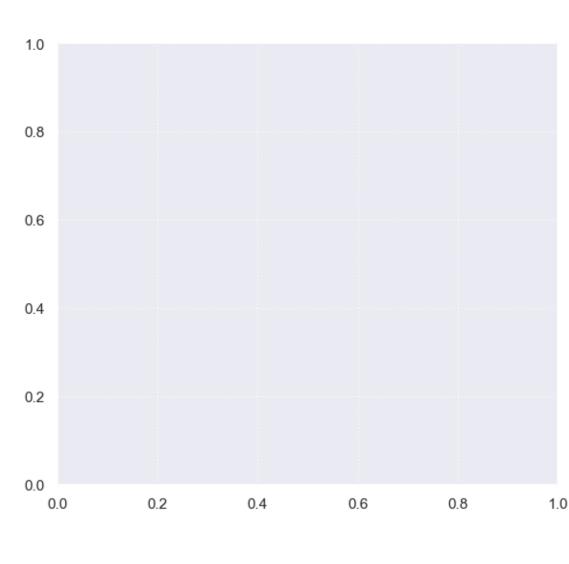
test: 0.0676

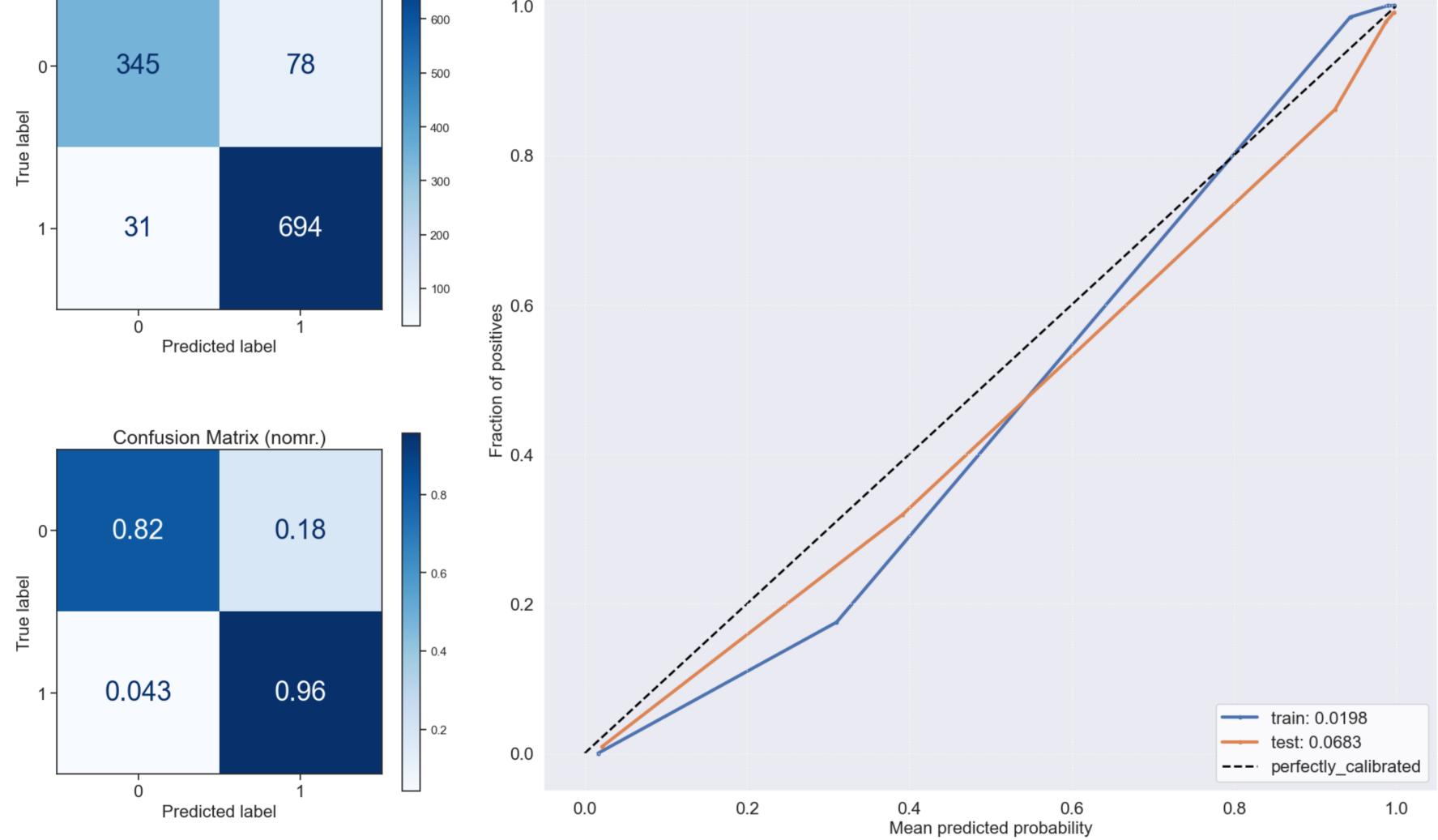
--- perfectly_calibrated

1.0

Protein: P56817 - metric: SoftVotingXGBoostEnsemble







			Active/Inactive pro	bability		
40	I	nactive zone			Active zone	type active (725) inactive (423) decoy (1143)
30						
Percent 05						
10						
0	0.0	0.2	0.4 Probability	0.6	0.8	1.0

Metric	Train	Test
AUC	0.9983	0.9659
accuracy	0.978	0.9051
balanced accuracy	0.9707	0.8864
recall_(sens)	0.9983	0.9572
specificity	0.9432	0.8156
precision	0.9679	0.899
f1_score	0.9779	0.9038