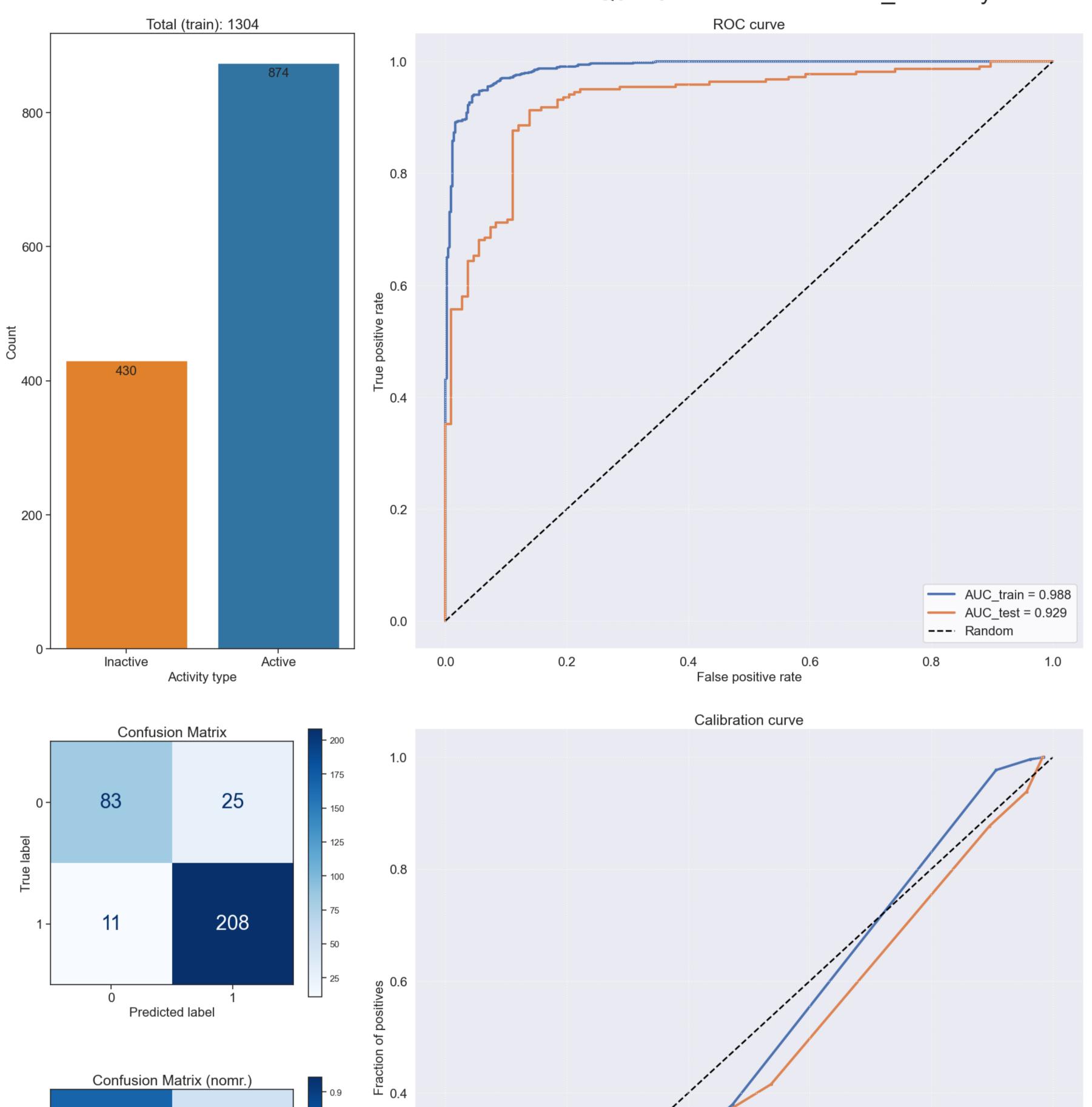
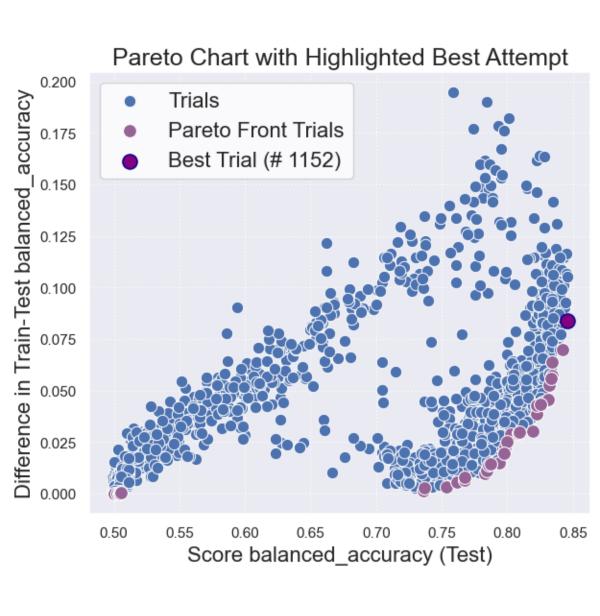
Protein: Q92731 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	2.4202
colsample_bytree	0.6394
gamma	0.7216
lambda	3.0648
learning_rate	0.0142
max_depth	10.0
min_child_weight	2.0
n_estimators	817.0
reg_alpha	0.0075
reg_lambda	0.8494
scale_pos_weight	1.1961
subsample	0.5872

	Act	tive/Inactive probabil	ity	
	Inactive zone		Active zone	type active (219) inactive (108) decoy (327)
15.0				
12.5				
10.0				
Dercent 7.5				
5.0				Ш
2.5				
0.0 —				
0.0	0.0 0.2 0.	4 0.0 Probability	6 0.8	1.0
i		Trobability		

- 0.8

- 0.7

- 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

0.23

0.95

0.77

0.05

True label

Metric	Train	Test
AUC	0.9877	0.9291
accuracy	0.9456	0.8899
balanced accuracy	0.931	0.8591
recall_(sens)	0.9737	0.9498
specificity	0.8884	0.7685
precision	0.9466	0.8927
f1_score	0.9451	0.8878

--- train: 0.0433

0.8

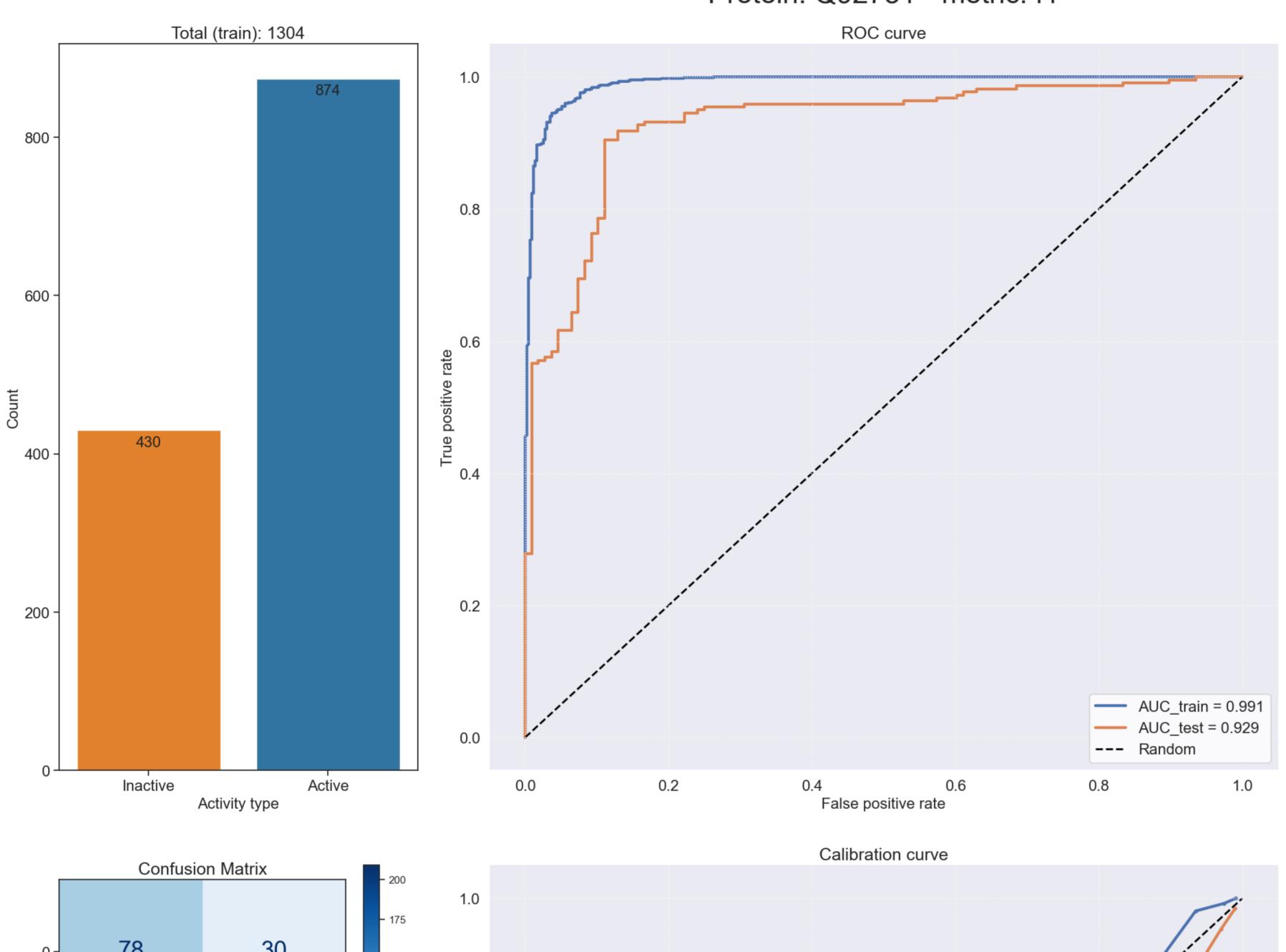
0.4 0.6 Mean predicted probability

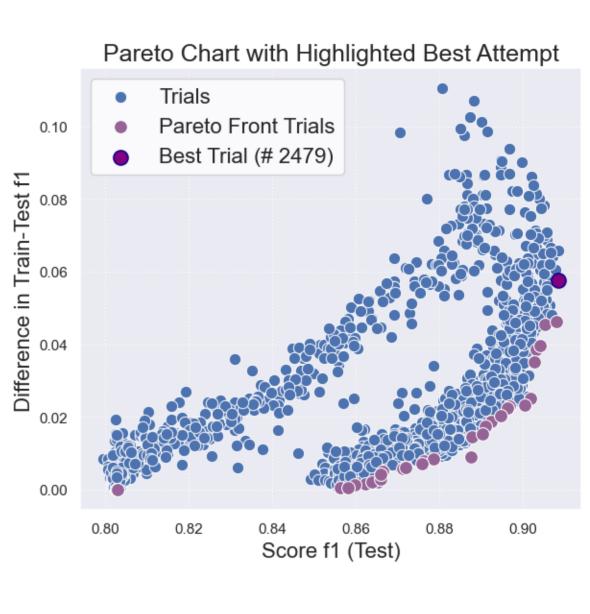
test: 0.0927

--- perfectly_calibrated

1.0

Protein: Q92731 - metric: f1





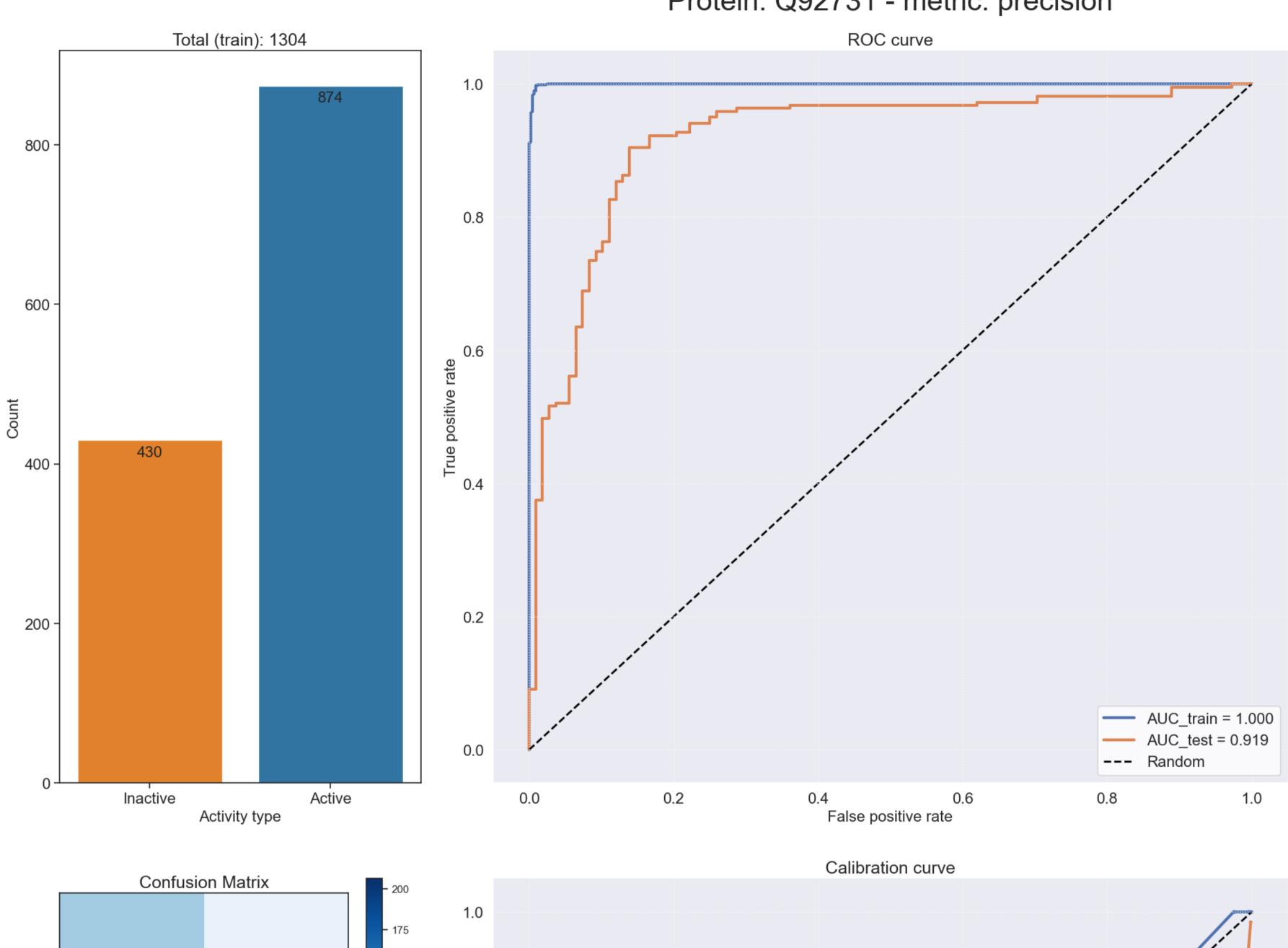
XGBoost Parameters	Value
alpha	1.5451
colsample_bytree	0.5655
gamma	0.3625
lambda	0.2511
learning_rate	0.0304
max_depth	8.0
min_child_weight	2.0
n_estimators	658.0
reg_alpha	0.7399
reg_lambda	0.3869
scale_pos_weight	1.7848
subsample	0.5312

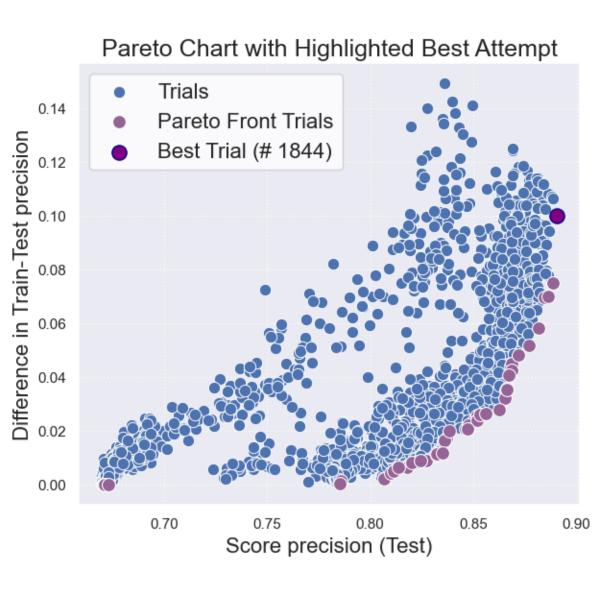
True label	78	30	- 200 - 175 - 150 - 125	1.0						
True	10	209	- 100 - 75 - 50 - 25	0.8				, or		
	0 Predicte	1 ∋d label		on of positives			, no			
	Confusion M	latrix (nomr.)		Fraction 6.0			and the second			
label O	0.72	0.28	- 0.8 - 0.6	0.2						
True label	0.046	0.95	- 0.4 - 0.2	0.0	para de la companya della companya d				train: 0.0	
	0 Predicte	1 ed label		0	0.0	0.2	0.4 Mean predicted	0.6 d probability	0.8	1.0

	Active/Inactive probability							
25	lna	ctive zone		Active zone	active (219) inactive (108) decoy (327)			
Percent 15 -								
10								
0 —					10			
	0.0	.2 0.4 Pr	0.6 obability	0.8	1.0			

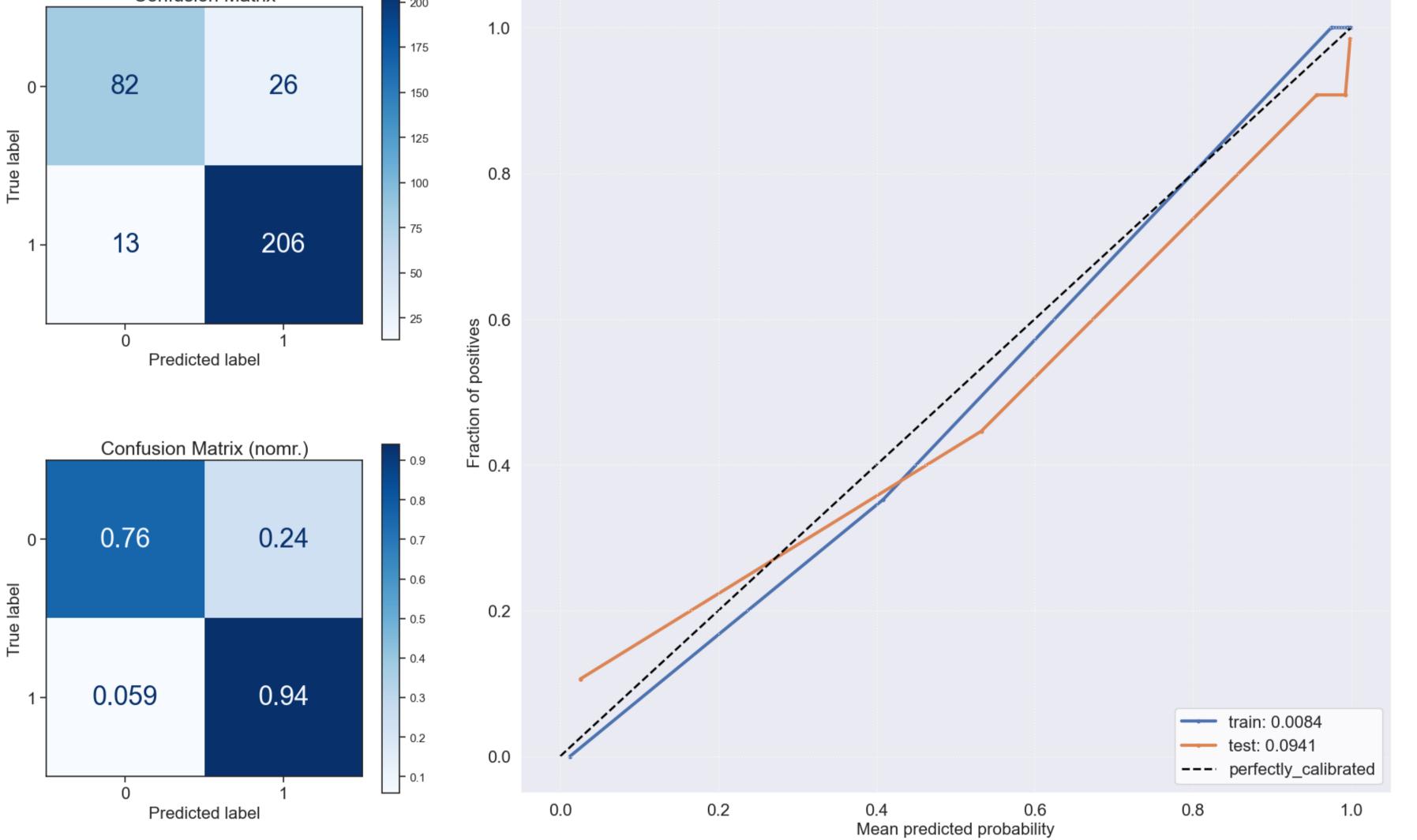
Metric	Train	Test
AUC	0.9908	0.9288
accuracy	0.9517	0.8777
balanced accuracy	0.9309	0.8383
recall_(sens)	0.992	0.9543
specificity	0.8698	0.7222
precision	0.9393	0.8745
f1_score	0.9509	0.8741

Protein: Q92731 - metric: precision



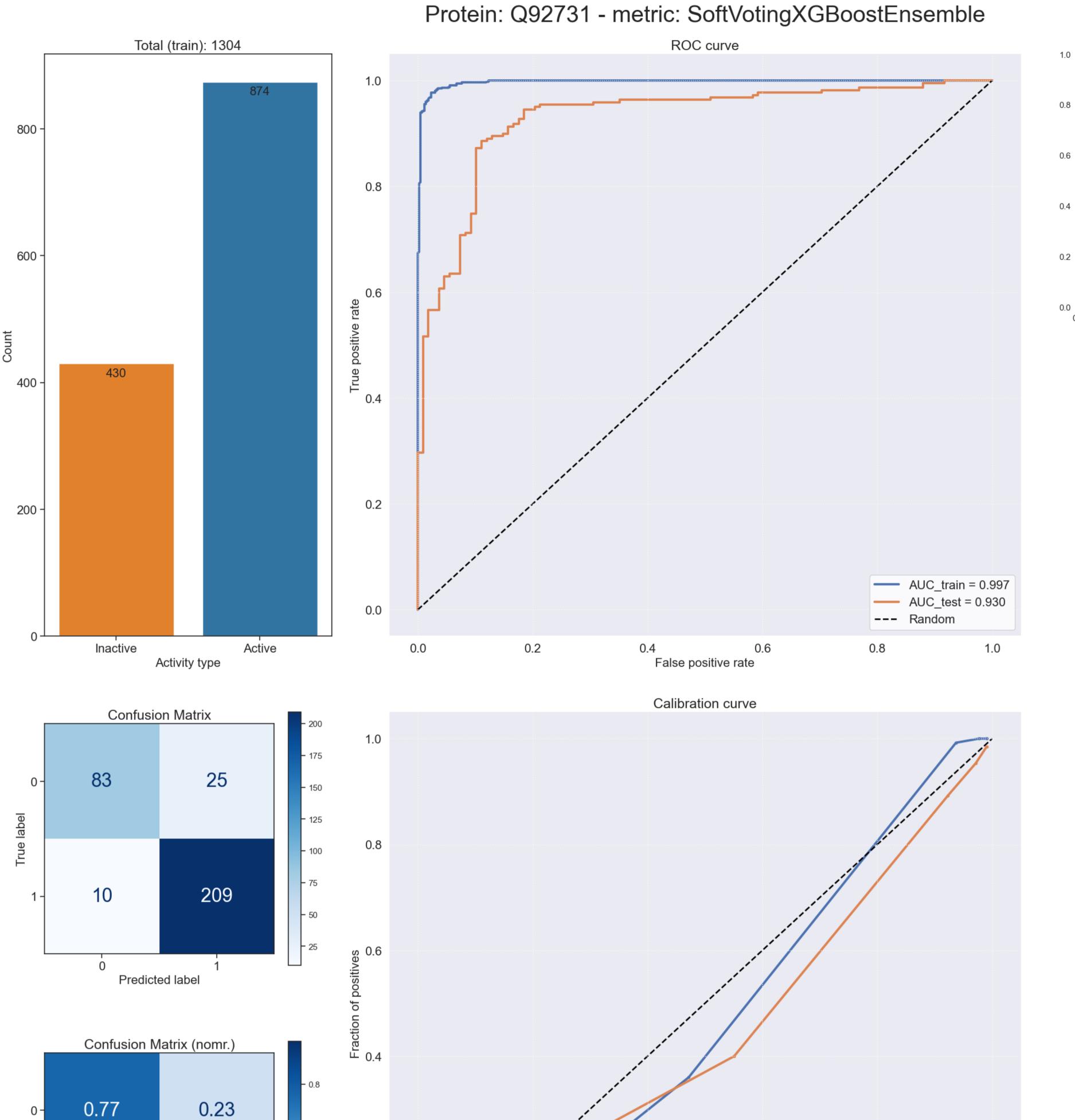


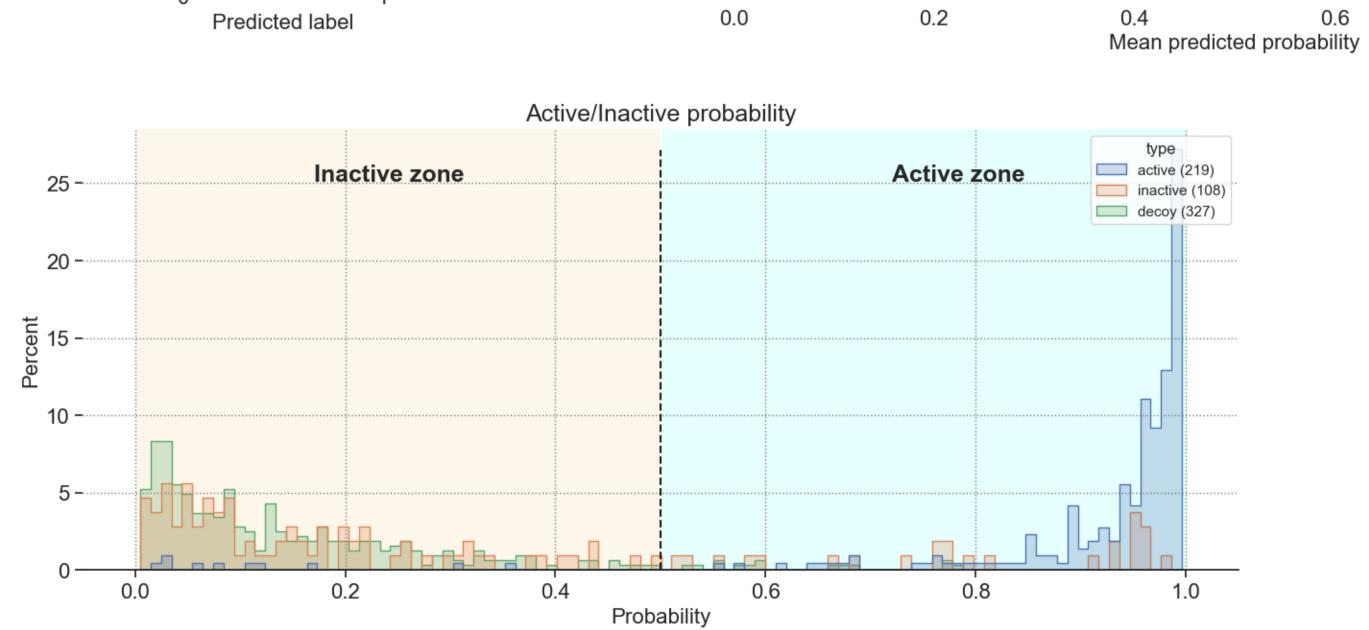
XGBoost Parameters	Value
alpha	4.7725
colsample_bytree	0.9572
gamma	0.3917
lambda	1.2964
learning_rate	0.4324
max_depth	9.0
min_child_weight	1.0
n_estimators	613.0
reg_alpha	0.1738
reg_lambda	0.0137
scale_pos_weight	1.5138
subsample	0.9005



	Active/Inactive probability								
50		tive zone			Active zone	active (219) inactive (108) decoy (327)			
40									
Percent 20									
10									
0 —			4	0.6					
	0.0 0.2	0.	4 Probabil	0.6 ity	0.8	1.0			

Metric	Train	Test
AUC	0.9996	0.9191
accuracy	0.9939	0.8807
balanced accuracy	0.9913	0.8499
recall_(sens)	0.9989	0.9406
specificity	0.9837	0.7593
precision	0.992	0.8879
f1_score	0.9939	0.8786





0.2

0.0

- 0.6

0.4

0.2

0.95

True label

0.046

Metric	Train	Test
AUC	0.9971	0.9297
accuracy	0.9716	0.893
balanced accuracy	0.9599	0.8614
recall_(sens)	0.9943	0.9543
specificity	0.9256	0.7685
precision	0.9645	0.8932
f1_score	0.9714	0.8907

0.2

0.6

0.0

Metric	Train	Test
AUC	0.9971	0.9297
accuracy	0.9716	0.893
balanced accuracy	0.9599	0.8614
recall_(sens)	0.9943	0.9543
specificity	0.9256	0.7685
precision	0.9645	0.8932
f1_score	0.9714	0.8907

--- train: 0.0261

--- test: 0.0909

0.8

--- perfectly_calibrated

1.0