Machine learning prediction of methionine and tryptophan photooxidation susceptibility

Jared A. Delmar,[1] Eugen Buehler,[2] Ashwin K. Chetty,[3] Agastya Das,[4] Guillermo Miro Quesada,[1] Jihong Wang,[1] and Xiaoyu Chen[1]

[1]Biopharmaceuticals Development, R&D, AstraZeneca, Gaithersburg, MD 20878, USA; [2]Data Sciences and AI, R&D, AstraZeneca, Gaithersburg, MD 20878, USA; [3]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA; [4]Khoury College of Computer Sciences, Northeastern University, Boston, MA 02115, USA

Photooxidation of methionine (Met) and tryptophan (Trp) residues is common and includes major degradation pathways that often pose a serious threat to the success of therapeutic proteins. Oxidation impacts all steps of protein production, manufacturing, and shelf life. Prediction of oxidation liability as early as possible in development is important because many more candidate drugs are discovered than can be tested experimentally. Undetected oxidation liabilities necessitate expensive and time-consuming remediation strategies in development and may lead to good drugs reaching patients slowly. Conversely, sites mischaracterized as oxidation liabilities could result in overengineering and lead to good drugs never reaching patients. To our knowledge, no predictive model for photooxidation of Met or Trp is currently available. We applied the random forest machine learning algorithm to in-house liquid chromatography-tandem mass spectrometry (LC-MS/MS) datasets (Met, n = 421; Trp, n = 342) of tryptic therapeutic protein peptides to create computational models for Met and Trp photooxidation. We show that our machine learning models predict Met and Trp photooxidation likelihood with 0.926 and 0.860 area under the curve (AUC), respectively, and Met photooxidation rate with a correlation coefficient (Q^2) of 0.511 and root-mean-square error (RMSE) of 10.9%. We further identify important physical, chemical, and formulation parameters that influence photooxidation. Improvement of biopharmaceutical liability predictions will result in better, more stable drugs, increasing development throughput, product quality, and likelihood of clinical success.

INTRODUCTION

Oxidation of methionine (Met) and tryptophan (Trp) residues are among the most common degradation pathways and affect all proteins.[1–3] In therapeutic proteins, oxidation impacts all production steps as well as the drug product throughout shelf life.[2,4,5] Often, oxidation is a serious threat to the success of therapeutic proteins, affecting both *in vitro* stability and *in vivo* biological function. Oxidation of Met and Trp residues has been demonstrated to negatively impact target affinity,[6–12] thermal stability,[13–16] biological activity,[7,9,17–23] serum half-life,[13,14,24–26] and immunogenicity.[27–33] Met oxidation is almost always a critical quality attribute in monoclonal

antibodies (mAbs) due to its impact on FcRn and FcγR binding, mediated by conserved heavy chain (HC) residues.[14,34–37] In many cases, oxidation of critical variable region residues will also necessitate a control strategy and monitoring during manufacturing and release. For example, a single Trp located in the HC complementarity determining region 3 (CDR3) of one humanized mAb was demonstrated to be singly responsible for its ultraviolet (UV) sensitivity, resulting in both loss of binding and loss of neutralization of its respiratory syncytial virus target.[7] Oxidation has also been observed to increase susceptibility to other degradation pathways, such as fragmentation and aggregation.[8,17,38–42] In another human immunoglobulin G1 (IgG1) mAb, photostress induced discoloration in the high-concentration liquid drug product, in addition to Trp oxidation in the light chain (LC) CDR3 and a concomitant loss of potency.[17] Met oxidation in particular has been shown to affect the function of diverse non-mAb therapeutic proteins.[9,19,43–45]

Although all 20 aa can be oxidized, including the protein backbone, observed oxidation rates span 3 orders of magnitude.[3,46] Practically, the most easily oxidized amino acids, and the amino acids of most concern for protein pharmaceuticals, are Met and Trp.[6,8,35,47,48] In the laboratory, accelerated oxidation of Met and Trp is typically achieved by chemical treatment with hydrogen peroxide (H_2O_2), 2,2′-azobis(2-amidinopropane) dihydrochloride (AAPH), or cool white light (CWL) and UV light irradiation.[6] However, while H_2O_2 and AAPH are useful for enriching oxidized species for further testing, they are not ideal stress conditions for assessing developability.[35] H_2O_2 treatment will preferentially oxidize Met and not Trp.[8,12,47] While AAPH treatment can promote oxidation of both Trp and Met, it may also introduce other modifications, such as covalent aggregation via dityrosine formation[10,49] and is not a relevant oxidizing agent to protein pharmaceutical manufacturing or storage conditions.[23] Alternatively, photooxidation is a known major contributor to oxidative degradation that affects both Met and Trp

Table 1. Predictors for methionine and tryptophan photooxidation machine learning models

Feature	Description
Fd	
Wd	closest approach atom-atom distance between M/W and nearest F/W/Y[4,55]
Yd	
SASA	
PSA	
SC_SASA	residue and side chain-only solvent accessibility
PSSA	
RefSASA	
Phi	
Psi	main and side-chain dihedral angles
Chi1	
Chi2	
Loop	
Helix	secondary structure element
Sheet	
Arginine	
Histidine	formulation free amino acid concentration in mM
Proline	
BondedStretch	
BondedBend	
BondedTorsion	bonded and non-bonded OPLS force field energy components[61]
BondedImpTor	
NonBondedInternal	
NonBondedInteraction	
pH	formulation pH
Sucrose	
Trehalose	formulation sugar concentration in mM
Polysorbate	formulation polysorbate concentration in %
Wexpscale	no. of surface exposed Trps and Mets per molecule, normalized to 150 kDa for non-mAbs
Mexpscale	

30 total features, or predictors, were used to inform the categorical and regression machine learning models to predict photooxidation likelihood and rate.

residues.[7,48,50] UV/CWL exposure is the only stress condition with International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidelines, and many commercial protein therapeutics carry warning labels to protect them from light.[17,50]

Early and accurate prediction of photooxidation as a development liability is important because many more candidate drugs are proposed than can be tested experimentally. "Latent" oxidation liabilities that are not dealt with as early as possible will require more expensive and time-consuming remediation strategies and could lead to good drugs reaching patients slowly. Use of oversimplified models that tend to overestimate oxidation risk is also problematic and will result

in overlooking or overengineering good drugs that, in turn, may never reach patients.

Many computational tools already exist to facilitate drug candidate screening, including advanced models based on machine learning.[51–56] However, oxidation models available to date were designed to predict only oxidation induced by H_2O_2[3,57,58] and AAPH chemical stress,[48,56,59] as well as *in vivo* oxidation that is often enzymatically driven.[55] Studies to date are limited to either Met[3,55–58] or Trp[48,59] residues.

In this study, we applied machine learning to liquid chromatography-tandem mass spectrometry (LC-MS/MS) datasets of therapeutic protein peptides containing Met (n = 421) and Trp (n = 342) to create accurate random forest[60] models for photooxidation. We show that our categorical models predict Met and Trp photooxidation likelihood ("yes" or "no") with a 0.926 and 0.860 mean area under the curve (AUC), respectively, determined by 5-fold cross-validation. In addition to Met photooxidation probability, we are able to accurately predict Met photooxidation rate by regression modeling, with correlation coefficient (Q2) of 0.511 and a root-mean-square error (RMSE) of 10.9%.

RESULTS

Feature selection

Observations gleaned from literature informed 30 features used to predict the photooxidation probability for each Met and Trp in our dataset and photooxidation rate for each Met (Table 1). The rationales for inclusion of each feature and its role in photooxidation, supported by literature, follow in the Discussion. Solvent accessibility was expressed as the total surface area of each residue in $Å^2$ (reference solvent-accessible surface area [RefSASA]), total exposed area of each residue in $Å^2$ (SASA), percent of surface area that is solvent exposed (percent solvent accessibility [PSA]), total exposed area of each side chain in $Å^2$ (side-chain solvent accessible surface area [SC_SASA]), and percent of side-chain surface area that is solvent exposed (percent side-chain solvent accessibility [PSSA]). The secondary structure of each residue was indicated by the binary parameters LOOP, SHEET, and HELIX. The side-chain and backbone conformations were accounted for by the dihedral angles phi, psi, chi1, and chi2. Sulfur-aromatic and aromatic-aromatic interactions were considered by the parameters Wd, Fd, and Yd which indicate the distance to the nearest Trp, phenylalanine (Phe), or tyrosine (Tyr) in Å, respectively.[55] Each component of the bonded and non-bonded Met or Trp residue energy, using the optimized potentials for liquid simulations (OPLS) force field,[61] was included in the parameters BondedStretch, BondedBend, BondedTorsion, BondedImpTor, NonBondedInternal, and NonBondedInteraction. The total numbers of solvent-exposed (percent solvent accessibility >10%) Met and Trp in each protein were defined as Mexpscale and Wexpscale, respectively. For non-mAb molecules, these two parameters were normalized by a scaling factor MW/150 kDa, where MW is the molecular weight of the non-mAb molecule in kDa and 150 kDa is the approximate mass of mAbs considered in this study. As each molecule in our study

Table 2. Confusion matrix for predictions made by the categorical machine learning model for predicting Met photooxidation probability on the independent holdout dataset

Prediction →		
Experiment ↓	Positive	Negative
Positive	3	0
Negative	0	11

Table 3. Statistics for predictions made by the categorical machine learning model for predicting Met photooxidation probability on the independent holdout dataset

Statistic	Met categorical model
Accuracy (%)	100.0
MCC	1.000
Precision (%)	100.0
Sensitivity (%)	100.0
Specificity (%)	100.0

was photostressed at the same mass concentration (10 mg/mL), normalizing the number of exposed Met or Trp per molecule by molecular weight results in a comparable number density of exposed Met or Trp in solution.

Formulation effects were taken into account by the predictors polysorbate (polysorbate concentration expressed in %), trehalose, sucrose, histidine (His), proline, and arginine (concentrations expressed in mM) and pH.

Because the predictive models developed here are most valuable during candidate selection, when the sequence information of a large number of potential drugs is known but little to no experimental data are available, we must be able to extract these parameters from the amino acid sequence. To determine the structural features, the 3D structure of each protein was first generated from the amino acid sequence by homology modeling.[62] Correlation between each feature and the experimental photooxidation measured by LC-MS/MS are shown in Tables S1 and S2.

Training and validation dataset construction

To satisfy the considerable data requirement of machine learning methods, we performed a side-by-side forced degradation study for 48 in-house molecules, including both mAbs and non-mAb therapeutic proteins. Photooxidation of each molecule was induced by exposure to CWL at 1.2 million lux h and UV light at 200 W h per square meter, per ICH guideline Q1B. Oxidation of Met and Trp residues was quantified by tryptic peptide mapping LC-MS/MS, and the experimental oxidation rate (%) was used to train the regression model for Met photooxidation rate. Of note, this dataset (Met, n = 421; Trp, n = 342) is much larger than other published datasets in the biopharmaceutical development field, but it is considered small in the broader field of machine learning.

To train categorical models for prediction of Met and Trp photooxidation probability, the experimental oxidation rates from tryptic peptide mapping were interpreted as "yes" or "no" based on a 5.0% threshold. For example, if we observed a 4.9% increase in oxidation of a certain Met after ICH light exposure, compared to the starting material, that site was not treated as photooxidized by the Met categorical model (class "no"). Alternatively, when we observed a 5.1% increase in oxidation of a certain Met, the model was trained to treat this site as oxidized (class "yes"). Oxidation abundance above this threshold is large enough to be considered a development liability and a significant change relative to the LC-MS/MS method variability.

All data available were used to train the Met regression model for prediction of photooxidation rate (n = 421). However, only one copy of each IgG-conserved Met and Trp was used to train each respective categorical model (Met, n = 235; Trp, n = 342). While more data are generally desirable and lead to more accurate machine learning models, trimming residues with highly conserved structure (in this case, mainly IgG1 Fc region residues) allowed us to reduce redundancy and bias of the models toward these residues, leading to more generalizable predictions. All three models were trained using stratified 5-fold cross-validation. Performance on independent holdout sets, comprised of Met and Trp variable region sites from four unseen mAbs not included in the training datasets, was also evaluated as confirmation of the performance on the cross-validation folds (Data S1, S2, S3, S4, and S4; Tables 2, 3, 4, and 5). The observed photooxidation frequency in the holdout dataset was consistent with the training set (Table 6). The distribution of protein format in each dataset is shown in Figure S1.

Unfortunately, attempts to predict Trp photooxidation rate were unsuccessful and resulted in poorly generalizable models, indicated by low Q2 (data not shown). It is likely that additional parameters are needed to describe Trp photooxidation that were not included in this study.

Machine learning models for prediction of Met and Trp photooxidation probability and Met photooxidation rate

Both the Met and Trp classification models as well as the Met regression model were built in R version 4.0.3 using the ranger[63] version 0.12.1 and caret[64] version 6.0.86 libraries.

A 5-fold stratified cross-validation strategy was used and hyperparameter tuning was performed during model training. The Met regression model for prediction of photooxidation rate achieved an average R^2 of 0.511 and 10.9% RMSE on the five cross-validation folds (Figure 1A). Categorical models achieved average cross-validation AUCs of 0.926 and 0.860 for Met and Trp, respectively (Figures 1B and 1C). Comparison lasso[65] regression models were built to incorporate feature selection, evaluate relative importance of features, and assess the advantage of random forest models for photooxidation prediction. For each prediction, the random forest models outperformed the lasso models on each fold (Figures 1A–1C). To determine whether our models were generalizable and able to accurately describe new

Table 4. Confusion matrix for predictions made by the categorical machine learning model for predicting Trp photooxidation probability on the independent holdout dataset

Prediction →		
Experiment ↓	Positive	Negative
Positive	3	1
Negative	2	21

Table 6. Contents of training and holdout datasets for each model

	No. of Met or Trp (Oxidized/Total)	
Model	Training Set	Holdout Set
Met categorical	86/235	3/14
Met regression	251/421	3/14
Trp categorical	67/342	4/27

data beyond the training sets, both Met and Trp classification models and the Met regression model were applied to independent holdout sets containing Met and Trp variable region sites from four mAbs not included in the training/validation splits. (Tables 2, 3, 4, and 5; Figure 2). Both Met and Trp categorical models performed with high accuracy on these independent holdout sets (Tables 3 and 5). The Met categorical model correctly classified 3 of 3 photooxidation liable sites and 11 of 11 non-liable sites, and the Trp categorical model correctly identified 3 of 4 photooxidation liable sites and 21 of 23 non-liable residues (Tables 2 and 4). The regression model for predicting Met photooxidation rate was also successful on its respective independent holdout set, predicting % oxidation after ICH photostress with Q2 of 0.567 and RMSE of 15.5% (Figure 2). Taken together, these results indicate that each model is capable of predicting site-specific Met and Trp oxidation probability, or, in the case of Met, oxidation rate, after UV and CWL stress.

The top predictors of Met photooxidation probability, Trp photooxidation probability, or Met photooxidation rate are shown in Figure 3. Of note, both distance to the nearest Phe (Fd) and solvent exposure (SASA, PSA, SC_SASA, or PSSA) appear among the top features used by all models. A detailed discussion of each feature and its importance to Met or Trp photooxidation appears in the following section.

DISCUSSION

Photooxidation of Met and Trp occurs as a result of either a type I or type II reaction.[66] Type I photooxidation involves photoinduced electron transfer by Trp to form radicals that react with ground state oxygen.[50,67] Other reactive oxygen species (ROS) can be formed as a byproduct of this process, including oxygen, hydroperoxyl, and hydroxyl radicals,[66] as well as H_2O_2, that readily oxidize Met.

Table 5. Statistics for predictions made by the categorical machine learning model for predicting Trp photooxidation probability on the independent holdout dataset.

Statistic	Met categorical model
Accuracy (%)	88.9
MCC	0.606
Precision (%)	60.0
Sensitivity (%)	75.0
Specificity (%)	91.3

While only Trp is a target for degradation in a type I process, both Trp and Met are susceptible to type II photooxidation (Figures 4 and 5).[66] In a type II photosensitization reaction, UV light absorbed by the protein is transferred to ground state oxygen, generating the reactive excited state singlet oxygen, $^1O_2$.[66] While Trp, Tyr, Phe, His, and cysteine (Cys) all have absorbance in the UV spectrum, other amino acids do not absorb significantly at wavelengths above 230 nm, including Met and the backbone.[68] In general, even the highly absorbing amino acids listed above are not efficient photosensitizers. Alternatively, the oxidized degradation products of Trp, such as kynurenine and n-formylkynurenine, are much more efficient.[68,69] Thus, extended UV exposure can quickly escalate photooxidation of both Met and Trp, as a result of interaction with $^1O_2$.[50,67]

Reaction of Met with $^1O_2$ first generates the persulfoxide intermediate (Figure 4). At acidic pH, the persulfoxide intermediate reacts with a second Met, forming two molecules of Met sulfoxide.[70–72] At pH above 9, the preferred pathway involves nucleophilic substitution on the sulfur atom, resulting in a single Met sulfoxide.[70,71] Although we do observe the degradation product Met sulfone after photostress, the pathway to transfer an additional oxygen to Met sulfoxide is not well characterized. Met sulfoxide- and Met sulfone-containing peptides were used to quantify Met photooxidation by LC-MS/MS for our training and holdout datasets.

Singlet oxygen reaction with Trp forms an unstable dioxetane intermediate that quickly decomposes via pyrrole ring cleavage to n-formylkynurenine (Figure 5).[73,74] Hydrolysis of n-formylkynurenine gives the degradation product kynurenine. While further oxidation of kynurenine to 3-hydroxykynurenine occurs rarely, this pathway is not well characterized. Alternatively, Trp can be directly oxidized by hydroxyl radicals to yield hydroxytryptophan (Figure 5). Kynurenine-, n-formylkynurenine-, 3-hydroxykynurenine-, and hydroxytryptophan-containing peptides were used to quantify Trp photooxidation by LC-MS/MS for our training and holdout datasets.

Solvent exposure has been cited as a prerequisite for oxidation,[49] and there exist one-parameter models for Met and Trp chemical oxidation susceptibility based solely on SASA.[3,48,57] However, a few groups have observed that additional structural features, besides solvent exposure, are needed to explain variability in oxidation rates.[10,29,49,75,76] Solvent exposure, captured by the features SASA, PSA, SC_SASA, and PSSA, were found to be among the top features in all photooxidation models presented herein (Figure 3) except for the lasso regression model to predict Met photooxidation rate. For categorical prediction of Met
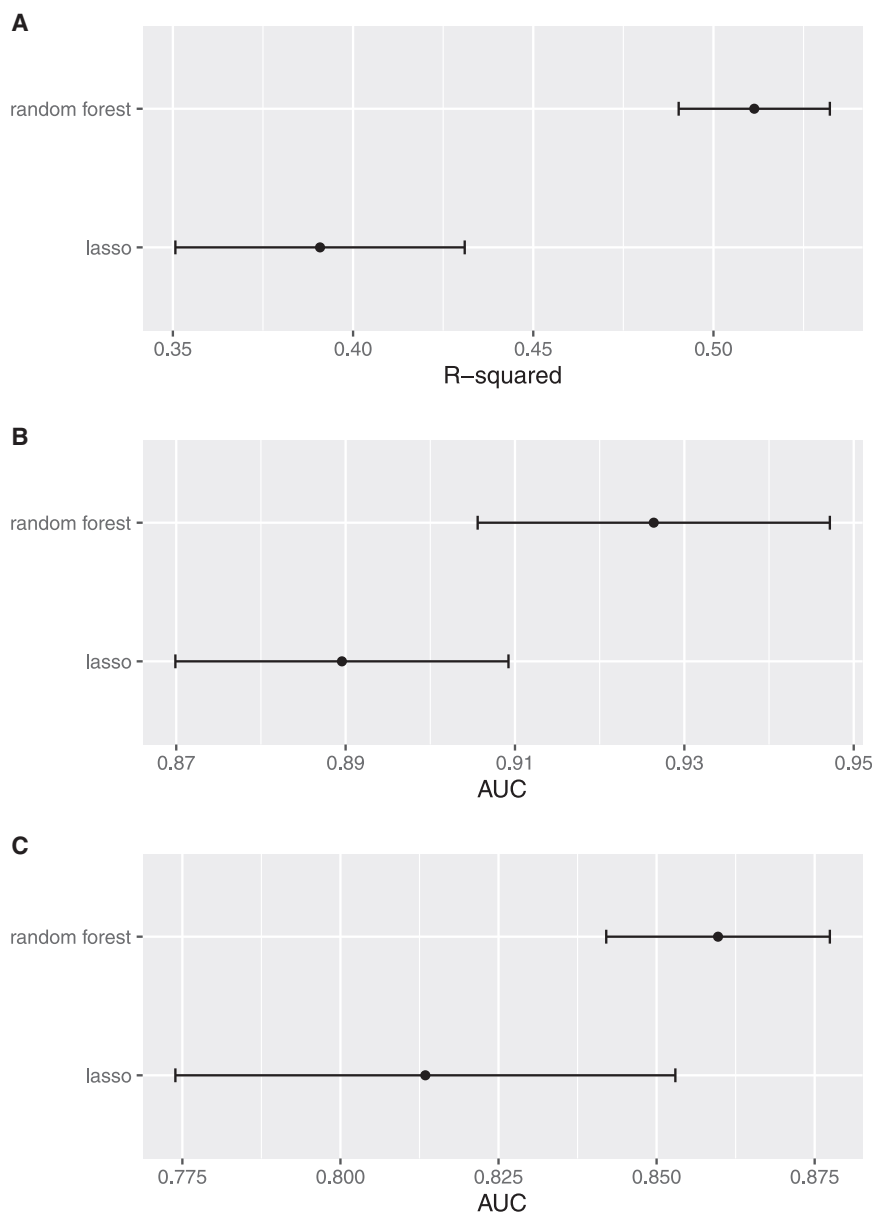
A

B

C

to protect proteins from oxidation.[78] In all predictive models of Met oxidation presented herein, Wd (distance to the nearest Trp) and Fd (distance to the nearest Phe) are among the most important parameters (Figure 3). The lasso coefficients corresponding to Wd and Fd are among the largest positive coefficients (Figure 3; Tables S1 and S2), supporting the observations of Aledo et al.[4,55] and Gray and Winkler.[78] Interestingly, the Met-Tyr interatomic distance (Yd) was among the least useful predictors in all Met models (Figure 3).

In addition to local interactions, distant photooxidation-prone residues in a protein may influence each other if ROS in solution are limited. Indeed, a scavenging role for Met has been proposed where surface-exposed Met residues in a structure act as antioxidants.[79] Alternatively, solvent-exposed Trp residues have been suggested to catalyze the photooxidation of neighboring residues.[80] While the total numbers of solvent-exposed Met or Trp residues, captured by the parameters Mexpscale and Wexpscale, were not found to be important for all Met or Trp categorical models in the present study, the random forest model for Met photooxidation rate relied on both Mexpscale and Wexpscale (Figure 3). Wexpscale was among the top features for our lasso regression model for Met photooxidation rate, and the corresponding lasso coefficient is positive (Tables S1 and S2), supporting the role of surface-exposed Trp in the generation of H_2O_2 proposed by Sreedhara et al.[80]

and Trp photooxidation, the lasso coefficients corresponding to solvent exposure are among the largest positive coefficients, indicating that increased solvent exposure increases likelihood of oxidation, as expected (Figure 3; Tables S1 and S2).

In one study, based on analysis of more than 2,000 Mets in 1,600 proteins subjected to H_2O_2 stress,[77] Aledo et al.[4,55] concluded that oxidation-prone Mets have different sequence environments than do oxidation-resistant Mets. Specifically, Mets in close proximity to aromatic side chains of Tyr, Trp, and Phe, indicative of sulfur-aromatic interactions, were found to be less prone to oxidize. Aromatic-aromatic interactions mediated by Tyr and Trp have also been proposed

Finally, the stability of biopharmaceuticals in liquid is well known to be affected by formulation conditions such as buffer composition, salt, excipients, and pH. Sugars and free amino acids can stabilize partially buried residues, limiting their reaction rates;[81–83] polysorbate surfactants are known to degrade to peroxides under heat and light stress;[84] and the Met oxidation pathway by $^1O_2$ is observed to be pH-dependent (Figure 1).[71] Both trehalose and polysorbate
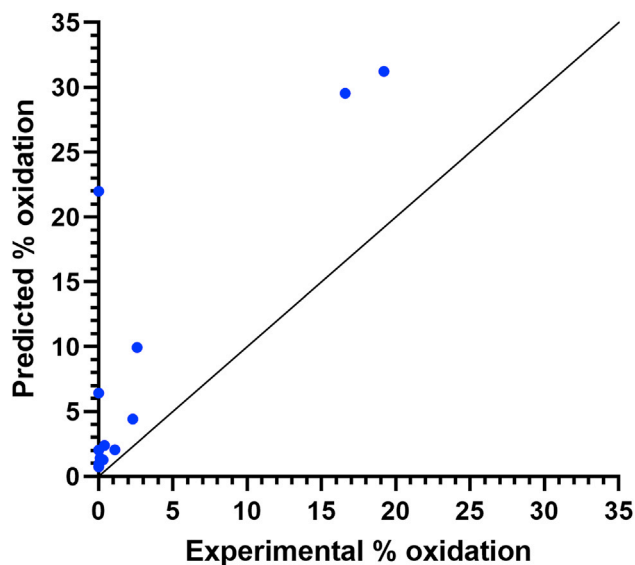
Figure 2. Regression machine learning model for predicting deamidation rate

Predicted Met oxidation abundance (%) was plotted versus the experimental measured oxidation abundance for the independent validation dataset. Individual Mets are plotted as blue circles, and the solid black line indicates where the predicted oxidation level equals the experimental oxidation level. Our regression model predicted the independent set with Q2 of 0.567 and RMSE of 15.5%.

concentrations were among the most important parameters for photooxidation rate prediction by lasso, and polysorbate is a top parameter in the photooxidation rate prediction by random forest (Figure 3). The lasso coefficients corresponding to polysorbate and trehalose concentrations were both negative, suggesting a protective effect of both formulation excipients (Figure 3; Tables S1 and S2).

No parameters describing the formulation buffer were found to be useful predictors of Trp photooxidation probability, for both lasso and random forest models, suggesting that Trp photooxidation and its effects cannot be mitigated by altering the buffer components considered in this study (Figure 3). However, the decreased accuracy compared to our Met photooxidation probability model and the failure to train an accurate model for Trp photooxidation rate may indicate that additional parameters are needed to describe Trp photooxidation that were not considered in this study.

Based on molecular dynamics simulations of oxidation liable Mets in granulocyte colony-stimulating factor (G-CSF), Chu et al.[5] speculated that protein conformation, including hydrogen bonding, may play important roles in oxidation induced by H_2O_2. Secondary structure features are utilized by both lasso models for Met photooxidation rate and Trp photooxidation probability (Figure 3).

Finally, the lasso model for Met photooxidation rate utilized the OPLS energy parameter BondedBend, indicating a negative correlation (Figure 3; Table S1). Of note, no other models, including the random

forest model for Met photooxidation rate, found high importance in the residue energy parameters. The relationship between residue energies and photooxidation liability of Met and Trp is not explored in the current literature.

The question remains whether oxidation susceptibility under AAPH or H_2O_2 stress is indicative of stability under normal storage and manufacturing conditions; especially H_2O_2, as peroxides are often used to clean filling lines for drug product manufacturing and can be generated from degradation of or included as impurities of formulation excipients.[79] However, in thermal and photostability studies, photooxidation is observed to preferentially target residues on the same antibody chain, while H_2O_2 treatment results in a random distribution.[35,85,86] This suggests a distinct pathway for photooxidation and the possibility that chemical oxidation studies may not represent oxidation that occurs under typical storage conditions.[35]

Conclusions

As of this writing, there are no published predictive models of Met or Trp photooxidation. Models available to date were only designed to predict oxidation induced by H_2O_2[3,58,59] and AAPH chemical stress,[49,57,60] as well as *in vivo* oxidation that is often enzymatically driven.[56]

We have trained models to predict photooxidation probability using peptide mapping data from both an antibody-variable region and non-antibody Met and Trp residues. These models predict photooxidation probability in Met and Trp with 0.926 and 0.860 AUC, respectively, evaluated by 5-fold cross-validation. In addition to photooxidation probability, we are able to accurately predict photooxidation rate for liable Met sites with Q^2 of 0.511 and RMSE of 10.9%. We have also evaluated our models on independent holdout datasets, comprising only mAb-variable region Met and Trp sites, indicating consistent performance. These models rely only on parameters available early in development when little to no experimental data have been generated for candidate sequences.

Currently, there is limited artificial intelligence (AI)-based support in protein therapeutic development and no available predictive models of photooxidation for proteins. It is our hope that with more data and increasingly accurate and interpretable models, a fundamental understanding of protein degradation, including oxidation, will be attained, leading to better and more stable drugs with increased development throughput and likelihood of clinical success.

MATERIALS AND METHODS

3D model building and parameter extraction

For AstraZeneca in-house molecules, full-length homology models were built using Schrödinger BioLuminate.[62] Briefly, the most similar crystal structure from the Protein Data Bank (PDB), by sequence, was first identified by basic local alignment search tool (BLAST).[87] This structure and an in-house constant region template were used as scaffolds for the full-length structure. The Protein Preparation Wizard tool was used to add hydrogens, assign bond orders, remove solvent
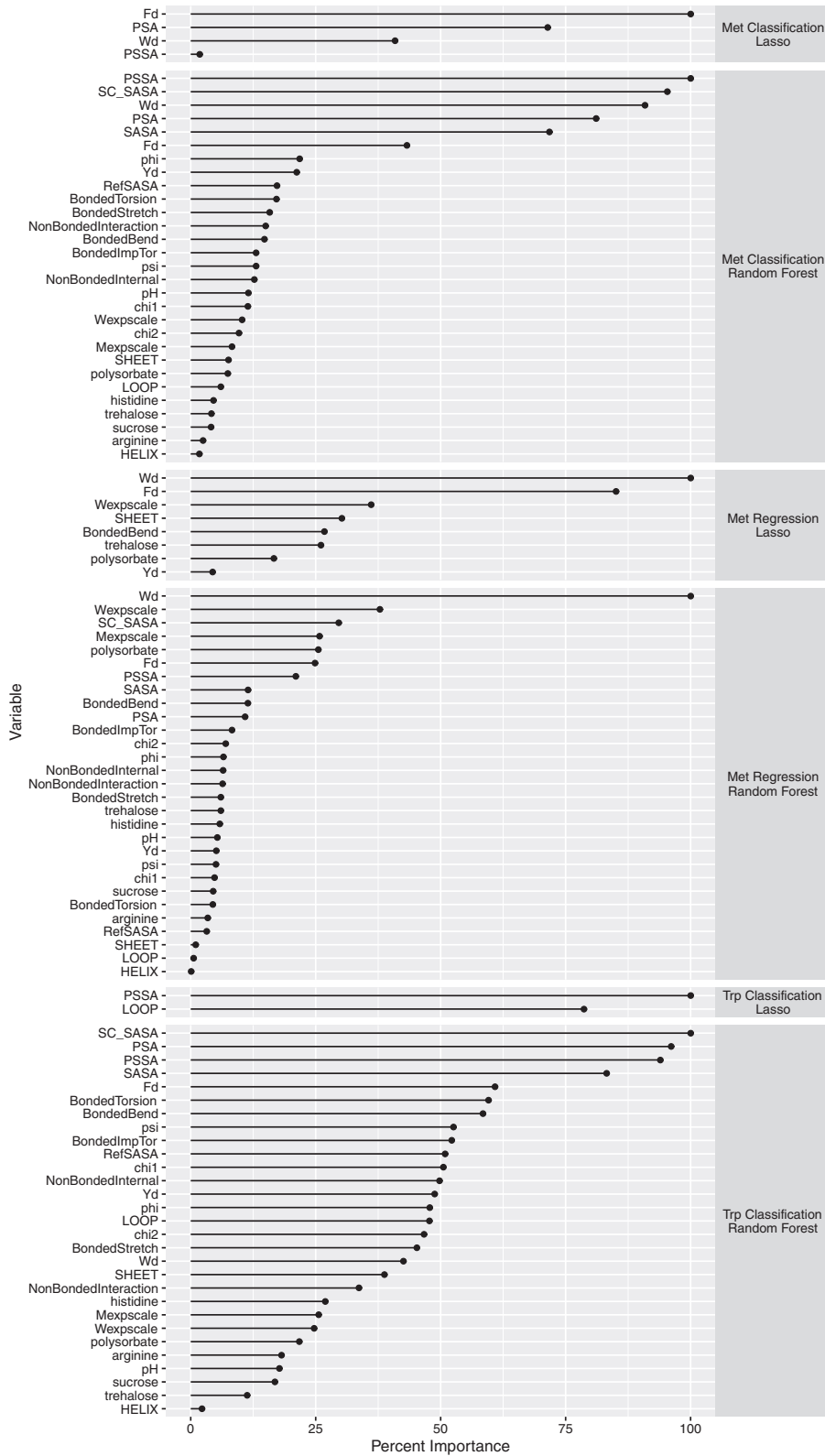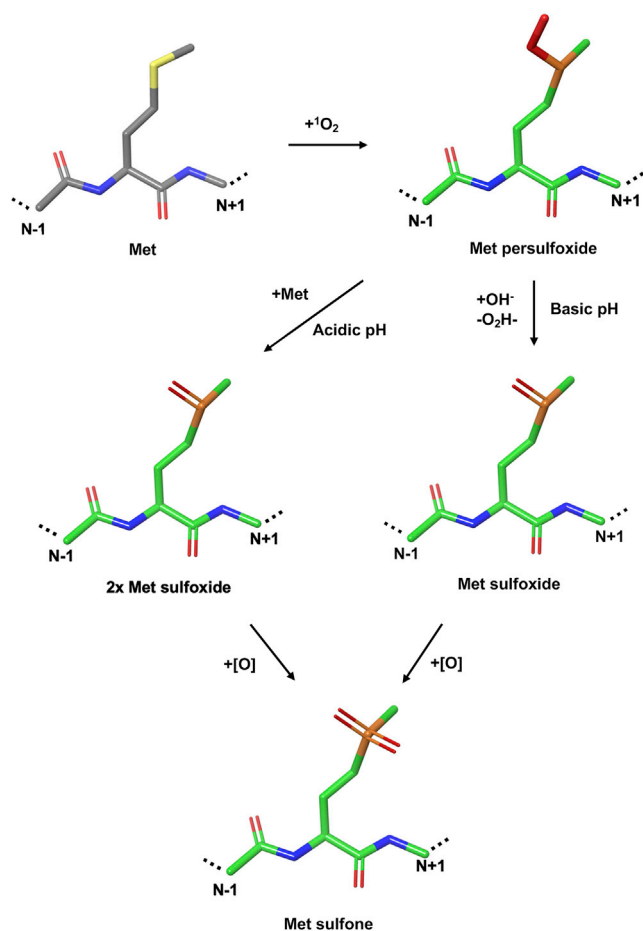
(legend on next page)

Figure 4. Met photooxidation pathway

Photooxidation of Met occurs, mainly, as a result of interaction with singlet oxygen, yielding the persulfoxide intermediate. At acidic pH, Met persulfoxide can interact with another Met, forming two molecules of Met sulfoxide. At basic pH, formation of Met sulfoxide occurs by reaction with a hydroxide ion, yielding a single Met sulfoxide (+16 Da). Further oxidation of Met sulfoxide results in Met sulfone (+32 Da). Residues are rendered as sticks with Met carbons and sulfurs colored gray and yellow, respectively, and Met photodegradation product carbons and sulfurs colored green and orange, respectively. Oxygen is colored red, and nitrogen is colored blue.

molecules, optimize H-bond assignments, and perform restrained energy minimization. Structural predictors of photooxidation were extracted from the 3D homology models within Schrödinger via Python and R[4,55] scripts.

Generation of photooxidized molecules

Each molecule was diluted to 250 μL at 10 mg/mL in a corresponding formulation buffer and aliquoted to LC/MS total recovery vials (Waters, Milford, MA, USA) and incubated at 25°C in a photostability chamber (Powers Scientific, Pipersville, PA, USA) to meet ICH guidelines for UV and CWL photoexposure. Reactants were stored at −80°C prior to analysis by LC-MS/MS.

LC-MS/MS tryptic peptide mapping

20-μL samples at 5 μg/μL were denatured by adding 200 μL of 8 M guanidine, 130 mM Tris, and 1 mM ethylenediaminetetraacetic acid (EDTA) pH 7.6 denaturing buffer. The samples were then reduced by the addition of 2 μL of 500 mM dithiothreitol. After incubation at 37°C for 30 min, samples were alkylated by the addition of 5 μL of 500 mM iodoacetamide and incubated at ambient temperature for 30 min in the dark. The reduced and alkylated samples were buffer exchanged into a solution containing 2 M urea and 100 mM Tris at pH 8.0 using an Amicon spin filter (EMD Millipore, Billerica, MA, USA; MW cutoff of 10 kDa); 5 μg of trypsin was then added to the spin filter and incubated at 37°C for 4 h. The digested samples were collected from the spin filters, and the digestion was quenched with trifluoroacetic acid.

Peptides produced by enzymatic digestion were eluted on an Acquity ultra performance LC system (Waters, Milford, MA, USA) equipped with an ethylene-bridged hybrid C18 reversed-phase column (1.7 μm, 2.1 mm, 150 mm) using a gradient of 0%–60% acetonitrile at a flow rate of 0.2 mL/min (total elution time of 76 min). Peptides separated on the column were identified by a UV detector and analyzed using an Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific). Peak identification was based on both the exact monoisotopic mass and the tandem mass spectrum of the target ion. Met and Trp oxidation quantitation was based on peak areas from the extracted ion chromatography of corresponding ions. Trp oxidation was quantified using the sum of kynurenine, *n*-formylkynurenine, 3-hydroxykynurenine, and hydroxytryptophan containing peptides. Met oxidation was quantified as the sum of Met sulfoxide and Met sulfone.

Random forest machine learning model construction

The best classification and regression models were random forest models built in R version 4.0.3 using the ranger[63] version 0.12.1 and caret[88] version 6.0.86 libraries. To compare to simpler, feature-limited models, lasso regressions were trained using the glmnet[89] version 4.1 library. The hyperparameters of all models were optimized by grid search using caret during training by stratified cross-validation.

For both Met and Trp classification models, 500 trees were generated with one variable tried at each split, producing cross-validation mean AUC of 0.926 and 0.860, respectively. The probability threshold at which we interpret the prediction as "yes" or "no" was 50%. Confusion matrices and variable importance plots were generated using the caret[64] library.

Figure 3. Met categorical, Met regression, and Trp categorical model feature importance

Relative importance of each parameter in the categorical model for predicting Met photooxidation probability, Trp photooxidation probability, and Met photooxidation rate was measured for the comparison lasso models and the random forest models. For lasso models, the importance is indicated by the magnitude of each coefficient. For random forest models, the importance was determined by the mean decrease in out-of-bag accuracy when that parameter was excluded from the model.
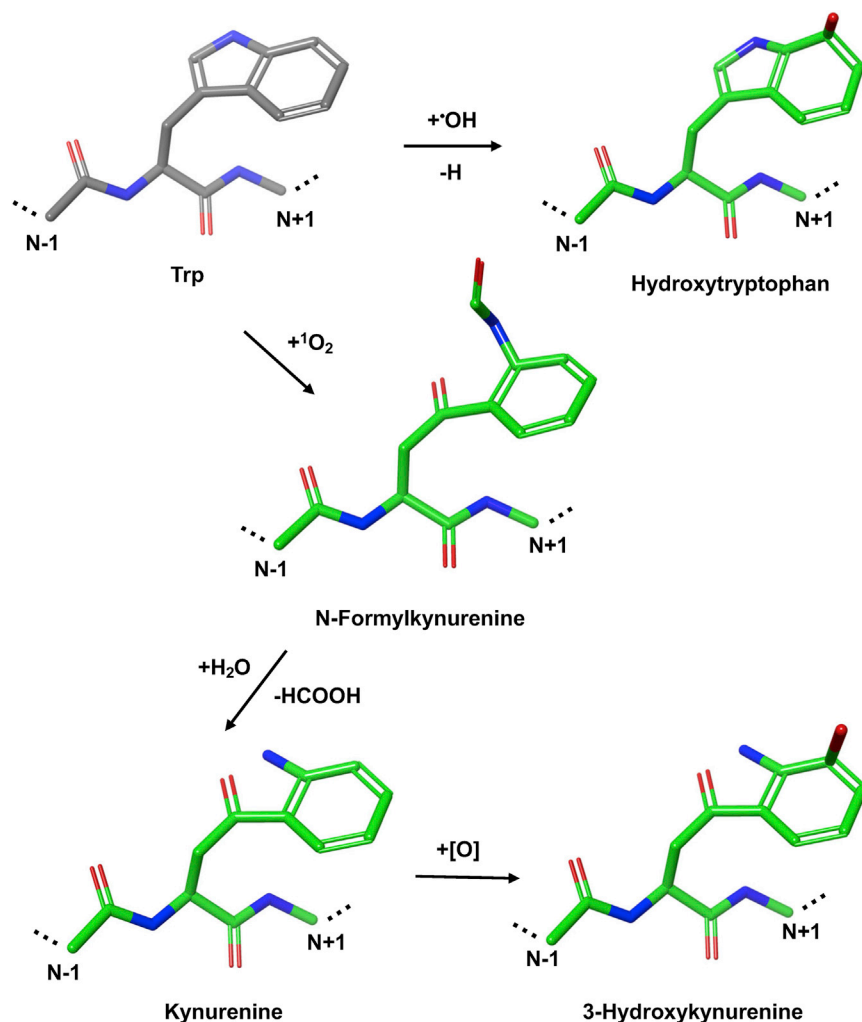
Figure 5. Trp photooxidation pathway

Photooxidation of Trp can occur by direct reaction with hydroxyl radicals, yielding hydroxytryptophan (+16 Da), or by reaction with singlet oxygen, yielding *n*-formylkynurenine (+32 Da) and kynurenine (+4 Da) degradation products. Further oxidation of kynurenine can produce 3-hydroxykynurenine (+20 Da). Residues are rendered as sticks with Trp carbons colored gray and Trp photodegradation product carbons colored green. Oxygen is colored red and nitrogen is colored blue.

The Met regression model was trained using 500 trees and 19 variables tried at each split. The cross-validated mean Q2 was 0.511. Q2 was calculated and variable importance plots were generated using the caret[64] library.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2021.03.023.

AUTHOR CONTRIBUTIONS

J.A.D. and J.W. conceptualized and designed the experiments; J.A.D., J.W., A.K.C., and A.D. performed the experiments and analyzed the data; J.A.D., E.B., A.K.C., and A.D. curated and interpreted the data and built and optimized the machine learning models; and J.A.D., E.B., A.K.C., A.D., G.M.Q., J.W., and X.C. wrote the manuscript.

DECLARATION OF INTERESTS

This work was supported by the global biologics R&D arm of AstraZeneca. J.A.D., E.B., G.M.Q., and X.C. are current employees of AstraZeneca and have stock and/or stock interests or options in AstraZeneca. The remaining authors declare no competing interests.

REFERENCES

1. Yan, Y., Wei, H., Fu, Y., Jusuf, S., Zeng, M., Ludwig, R., Krystek, S.R., Jr., Chen, G., Tao, L., and Das, T.K. (2016). Isomerization and oxidation in the complementarity-determining regions of a monoclonal antibody: A study of the modification-structure-function correlations by hydrogen-deuterium exchange mass spectrometry. Anal. Chem. 88, 2041–2050.

2. Li, S., Schöneich, C., and Borchardt, R.T. (1995). Chemical instability of protein pharmaceuticals: Mechanisms of oxidation and strategies for stabilization. Biotechnol. Bioeng. 48, 490–500.

3. Yang, R., Jain, T., Lynaugh, H., Nobrega, R.P., Lu, X., Boland, T., Burnina, I., Sun, T., Caffry, I., Brown, M., et al. (2017). Rapid assessment of oxidation via middle-down LCMS correlates with methionine side-chain solvent-accessible surface area for 121 clinical stage monoclonal antibodies. MAbs 9, 646–653.

4. Aledo, J.C., Cantón, F.R., and Veredas, F.J. (2015). Sulphur atoms from methionines interacting with aromatic residues are less prone to oxidation. Sci. Rep. 5, 16955.

5. Chu, J.W., Yin, J., Wang, D.I., and Trout, B.L. (2004). Molecular dynamics simulations and oxidation rates of methionine residues of granulocyte colony-stimulating factor at different pH values. Biochemistry 43, 1019–1029.

6. Dashivets, T., Stracke, J., Dengl, S., Knaupp, A., Pollmann, J., Buchner, J., and Schlothauer, T. (2016). Oxidation in the complementarity-determining regions differentially influences the properties of therapeutic antibodies. MAbs 8, 1525–1535.

7. Wei, Z., Feng, J., Lin, H.Y., Mullapudi, S., Bishop, E., Tous, G.I., Casas-Finet, J., Hakki, F., Strouse, R., and Schenerman, M.A. (2007). Identification of a single tryptophan residue as critical for binding activity in a humanized monoclonal antibody against respiratory syncytial virus. Anal. Chem. 79, 2797–2805.

8. Hensel, M., Steurer, R., Fichtl, J., Elger, C., Wedekind, F., Petzold, A., Schlothauer, T., Molhoj, M., Reusch, D., and Bulau, P. (2011). Identification of potential sites for tryptophan oxidation in recombinant antibodies using tert-butylhydroperoxide and quantitative LC-MS. PLoS ONE 6, e17708.

9. Teh, L.C., Murphy, L.J., Huq, N.L., Surus, A.S., Friesen, H.G., Lazarus, L., and Chapman, G.E. (1987). Methionine oxidation in human growth hormone and human chorionic somatomammotropin. Effects on receptor binding and biological activities. J. Biol. Chem. 262, 6472–6477.

10. Torosantucci, R., Schöneich, C., and Jiskoot, W. (2014). Oxidation of therapeutic proteins and peptides: Structural and biological consequences. Pharm. Res. 31, 541–553.

11. Haberger, M., Bomans, K., Diepold, K., Hook, M., Gassner, J., Schlothauer, T., Zwick, A., Spick, C., Kepert, J.F., Hienz, B., et al. (2014). Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: Susceptibility vs. functionality of critical quality attributes. MAbs 6, 327–339.

12. Stroop, S.D., Conca, D.M., Lundgard, R.P., Renz, M.E., Peabody, L.M., and Leigh, S.D. (2011). Photosensitizers form in histidine buffer and mediate the photodegradation of a monoclonal antibody. J. Pharm. Sci. 100, 5142–5155.

13. Gao, X., Ji, J.A., Veeravalli, K., Wang, Y.J., Zhang, T., Mcgreevy, W., Zheng, K., Kelley, R.F., Laird, M.W., Liu, J., and Cromwell, M. (2015). Effect of individual Fc methionine oxidation on FcRn binding: Met252 oxidation impairs FcRn binding more profoundly than Met428 oxidation. J. Pharm. Sci. 104, 368–377.

14. Bertolotti-Ciarlet, A., Wang, W., Lownes, R., Pristatsky, P., Fang, Y., McKelvey, T., Li, Y., Li, Y., Drummond, J., Prueksaritanont, T., and Vlasak, J. (2009). Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fcγ receptors. Mol. Immunol. 46, 1878–1882.

15. Liu, D., Ren, D., Huang, H., Dankberg, J., Rosenfeld, R., Cocco, M.J., Li, L., Brems, D.N., and Remmele, R.L., Jr. (2008). Structure and stability changes of human IgG1 Fc as a consequence of methionine oxidation. Biochemistry 47, 5088–5100.

16. Mulinacci, F., Capelle, M.A., Gurny, R., Drake, A.F., and Arvinte, T. (2011). Stability of human growth hormone: influence of methionine oxidation on thermal folding. J. Pharm. Sci. 100, 451–463.

17. Qi, P., Volkin, D.B., Zhao, H., Nedved, M.L., Hughes, R., Bass, R., Yi, S.C., Panek, M.E., Wang, D., Dalmonte, P., and Bond, M.D. (2009). Characterization of the photodegradation of a human IgG1 monoclonal antibody formulated as a high-concentration liquid dosage form. J. Pharm. Sci. 98, 3117–3130.

18. Reubsaet, J.L., Beijnen, J.H., Bult, A., Hop, E., Scholten, S.D., Teeuwsen, J., and Underberg, W.J. (1998). Oxidation of recombinant methionyl human granulocyte colony-stimulating factor. J. Pharm. Biomed. Anal. 17, 283–289.

19. Lu, H.S., Fausset, P.R., Narhi, L.O., Horan, T., Shinagawa, K., Shimamoto, G., and Boone, T.C. (1999). Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony-stimulating factor: effect on stability and biological activity. Arch. Biochem. Biophys. 362, 1–11.

20. Larocque, L., Bliu, A., Xu, R., Diress, A., Wang, J., Lin, R., He, R., Girard, M., and Li, X. (2011). Bioactivity determination of native and variant forms of therapeutic interferons. J. Biomed. Biotechnol. 2011, 174615.

21. Zull, J.E., Smith, S.K., and Wiltshire, R. (1990). Effect of methionine oxidation and deletion of amino-terminal residues on the conformation of parathyroid hormone. Circular dichroism studies. J. Biol. Chem. 265, 5671–5676.

22. Nabuchi, Y., Fujiwara, E., Ueno, K., Kuboniwa, H., Asoh, Y., and Ushio, H. (1995). Oxidation of recombinant human parathyroid hormone: Effect of oxidized position on the biological activity. Pharm. Res. 12, 2049–2052.

23. Barnett, G.V., Balakrishnan, G., Chennamsetty, N., Hoffman, L., Bongers, J., Tao, L., Huang, Y., Slaney, T., Das, T.K., Leone, A., and Kar, S.R. (2019). Probing the tryptophan environment in therapeutic proteins: implications for higher order structure on tryptophan oxidation. J. Pharm. Sci. 108, 1944–1952.

24. Wang, W., Vlasak, J., Li, Y., Pristatsky, P., Fang, Y., Pittman, T., Roman, J., Wang, Y., Prueksaritanont, T., and Ionescu, R. (2011). Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. Mol. Immunol. 48, 860–866.

25. Wang, S., Ionescu, R., Peekhaus, N., Leung, J.Y., Ha, S., and Vlasak, J. (2010). Separation of post-translational modifications in monoclonal antibodies by exploiting subtle conformational changes under mildly acidic conditions. J. Chromatogr. A 1217, 6496–6502.

26. Pan, H., Chen, K., Chu, L., Kinderman, F., Apostol, I., and Huang, G. (2009). Methionine oxidation in human IgG2 Fc decreases binding affinities to protein A and FcRn. Protein Sci. 18, 424–433.

27. Hermeling, S., Crommelin, D.J., Schellekens, H., and Jiskoot, W. (2004). Structure-immunogenicity relationships of therapeutic proteins. Pharm. Res. 21, 897–903.

28. Filipe, V., Jiskoot, W., Basmeleh, A.H., Halim, A., Schellekens, H., and Brinks, V. (2012). Immunogenicity of different stressed IgG monoclonal antibody formulations in immune tolerant transgenic mice. MAbs 4, 740–752.

29. Torosantucci, R., Mozziconacci, O., Sharov, V., Schöneich, C., and Jiskoot, W. (2012). Chemical modifications in aggregates of recombinant human insulin induced by metal-catalyzed oxidation: Covalent cross-linking via Michael addition to tyrosine oxidation products. Pharm. Res. 29, 2276–2293.

30. Torosantucci, R., Kukrer, B., Mero, A., Van Winsen, M., Tantipolphan, R., and Jiskoot, W. (2011). Plain and mono-pegylated recombinant human insulin exhibit similar stress-induced aggregation profiles. J. Pharm. Sci. 100, 2574–2585.

31. Ryff, J.C. (1997). Clinical investigation of the immunogenicity of interferon-alpha 2a. J. Interferon Cytokine Res. 17 (Suppl 1), S29–S33.

32. Hochuli, E. (1997). Interferon immunogenicity: technical evaluation of interferon-alpha 2a. J. Interferon Cytokine Res. 17 (Suppl 1), S15–S21.

33. Hermeling, S., Aranha, L., Damen, J.M., Slijper, M., Schellekens, H., Crommelin, D.J., and Jiskoot, W. (2005). Structural characterization and immunogenicity in wild-type and immune tolerant mice of degraded recombinant human interferon alpha2b. Pharm. Res. 22, 1997–2006.

34. Goetze, A.M., Schenauer, M.R., and Flynn, G.C. (2010). Assessing monoclonal antibody product quality attribute criticality through clinical studies. MAbs 2, 500–507.

35. Nowak, C., K Cheung, J., M Dellatore, S., Katiyar, A., Bhat, R., Sun, J., Ponniah, G., Neill, A., Mason, B., Beck, A., and Liu, H. (2017). Forced degradation of recombinant monoclonal antibodies: A practical guide. MAbs 9, 1217–1230.

36. Cymer, F., Thomann, M., Wegele, H., Avenal, C., Schlothauer, T., Gygax, D., and Beck, H. (2017). Oxidation of M252 but not M428 in hu-IgG1 is responsible for decreased binding to and activation of hu-FcγRIIa (His131). Biologicals 50, 125–128.

37. Eon-Duval, A., Broly, H., and Gleixner, R. (2012). Quality attributes of recombinant therapeutic proteins: An assessment of impact on safety and efficacy as part of a quality by design development approach. Biotechnol. Prog. 28, 608–622.

38. Mulinacci, F., Poirier, E., Capelle, M.A., Gurny, R., and Arvinte, T. (2013). Influence of methionine oxidation on the aggregation of recombinant human growth hormone. Eur. J. Pharm. Biopharm. *85*, 42–52.

39. Dion, M.Z., Leiske, D., Sharma, V.K., Zuch de Zafra, C.L., and Salisbury, C.M. (2018). Mitigation of oxidation in therapeutic antibody formulations: A biochemical efficacy and safety evaluation of *N*-acetyl-tryptophan and L-methionine. Pharm. Res. *35*, 222.

40. Luo, Q., Joubert, M.K., Stevenson, R., Ketchem, R.R., Narhi, L.O., and Wypych, J. (2011). Chemical modifications in therapeutic protein aggregates generated under different stress conditions. J. Biol. Chem. *286*, 25134–25144.

41. Joubert, M.K., Luo, Q., Nashed-Samuel, Y., Wypych, J., and Narhi, L.O. (2011). Classification and characterization of therapeutic antibody aggregates. J. Biol. Chem. *286*, 25118–25133.

42. Steinmann, D., Ji, J.A., Wang, Y.J., and Schöneich, C. (2012). Oxidation of human growth hormone by oxygen-centered radicals: Formation of Leu-101 hydroperoxide and Tyr-103 oxidation products. Mol. Pharm. *9*, 803–814.

43. Thirumangalathu, R., Krishnan, S., Bondarenko, P., Speed-Ricci, M., Randolph, T.W., Carpenter, J.F., and Brems, D.N. (2007). Oxidation of methionine residues in recombinant human interleukin-1 receptor antagonist: Implications of conformational stability on protein oxidation kinetics. Biochemistry *46*, 6213–6224.

44. Lam, X.M., Lai, W.G., Chan, E.K., Ling, V., and Hsu, C.C. (2011). Site-specific tryptophan oxidation induced by autocatalytic reaction of polysorbate 20 in protein formulation. Pharm. Res. *28*, 2543–2555.

45. Matheson, N.R., and Travis, J. (1985). Differential effects of oxidizing agents on human plasma alpha 1-proteinase inhibitor and human neutrophil myeloperoxidase. Biochemistry *24*, 1941–1945.

46. Davies, M.J. (2005). The oxidative environment and protein damage. Biochim. Biophys. Acta *1703*, 93–109.

47. Vogt, W. (1995). Oxidation of methionyl residues in proteins: Tools, targets, and reversal. Free Radic. Biol. Med. *18*, 93–105.

48. Pavon, J.A., Xiao, L., Li, X., Zhao, J., Aldredge, D., Dank, E., Fridman, A., and Liu, Y.H. (2019). Selective tryptophan oxidation of monoclonal antibodies: Oxidative stress and modeling prediction. Anal. Chem. *91*, 2192–2200.

49. Jarasch, A., Koll, H., Regula, J.T., Bader, M., Papadimitriou, A., and Kettenberger, H. (2015). Developability assessment during the selection of novel therapeutic antibodies. J. Pharm. Sci. *104*, 1885–1898.

50. Kerwin, B.A., and Remmele, R.L., Jr. (2007). Protect from light: Photodegradation and protein biologics. J. Pharm. Sci. *96*, 1468–1479.

51. Delmar, J.A., Wang, J., Choi, S.W., Martins, J.A., and Mikhail, J.P. (2019). Machine learning enables accurate prediction of asparagine deamidation probability and rate. Mol. Ther. Methods Clin. Dev. *15*, 264–274.

52. Hebditch, M., and Warwicker, J. (2019). Charge and hydrophobicity are key features in sequence-trained machine learning models for predicting the biophysical properties of clinical-stage antibodies. PeerJ *7*, e8199.

53. Obrezanova, O., Arnell, A., de la Cuesta, R.G., Berthelot, M.E., Gallagher, T.R., Zurdo, J., and Stallwood, Y. (2015). Aggregation risk prediction for antibodies and its application to biotherapeutic development. MAbs *7*, 352–363.

54. Jia, L., and Sun, Y. (2017). Protein asparagine deamidation prediction based on structures with machine learning methods. PLoS ONE *12*, e0181347.

55. Aledo, J.C., Cantón, F.R., and Veredas, F.J. (2017). A machine learning approach for predicting methionine oxidation sites. BMC Bioinformatics *18*, 430.

56. Sankar, K., Hoi, K.H., Yin, Y., Ramachandran, P., Andersen, N., Hilderbrand, A., McDonald, P., Spiess, C., and Zhang, Q. (2018). Prediction of methionine oxidation risk in monoclonal antibodies using a machine learning method. MAbs *10*, 1281–1290.

57. Chennamsetty, N., Quan, Y., Nashine, V., Sadineni, V., Lyngberg, O., and Krystek, S. (2015). Modeling the oxidation of methionine residues by peroxides in proteins. J. Pharm. Sci. *104*, 1246–1255.

58. Agrawal, N.J., Dykstra, A., Yang, J., Yue, H., Nguyen, X., Kolvenbach, C., and Angell, N. (2018). Prediction of the hydrogen peroxide-induced methionine oxidation propensity in monoclonal antibodies. J. Pharm. Sci. *107*, 1282–1289.

59. Sharma, V.K., Patapoff, T.W., Kabakoff, B., Pai, S., Hilario, E., Zhang, B., Li, C., Borisov, O., Kelley, R.F., Chorny, I., et al. (2014). In silico selection of therapeutic antibodies for development: viscosity, clearance, and chemical stability. Proc. Natl. Acad. Sci. USA *111*, 18601–18606.

60. Breiman, L. (2001). Random forests. Mach. Learn. *45*, 5–32.

61. Jorgensen, W.L., and Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. Soc. *110*, 1657–1666.

62. Zhu, K., Day, T., Warshaviak, D., Murrett, C., Friesner, R., and Pearlman, D. (2014). Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. Proteins *82*, 1646–1655.

63. Wright, M.N., and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. *77*, 10.18637/jss.v077.i01.

64. Kuhn, M. (2008). Building predictive models in R using the caret package. J. Stat. Softw. *28*, 1–26, 10.18637/jss.v028.i05.

65. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. R. Stat. Soc. B *58*, 267–288.

66. Baptista, M.S., Cadet, J., Di Mascio, P., Ghogare, A.A., Greer, A., Hamblin, M.R., Lorente, C., Nunez, S.C., Ribeiro, M.S., Thomas, A.H., et al. (2017). Type I and type II photosensitized oxidation reactions: Guidelines and mechanistic pathways. Photochem. Photobiol. *93*, 912–919.

67. Pattison, D.I., Rahmanto, A.S., and Davies, M.J. (2012). Photo-oxidation of proteins. Photochem. Photobiol. Sci. *11*, 38–53.

68. Davies, M.J. (2003). Singlet oxygen-mediated damage to proteins and its consequences. Biochem. Biophys. Res. Commun. *305*, 761–770.

69. Redmond, R.W., and Gamlin, J.N. (1999). A compilation of singlet oxygen yields from biologically relevant molecules. Photochem. Photobiol. *70*, 391–475.

70. Grassi, L., and Cabrele, C. (2019). Susceptibility of protein therapeutics to spontaneous chemical modifications by oxidation, cyclization, and elimination reactions. Amino Acids *51*, 1409–1431.

71. Di Mascio, P., Martinez, G.R., Miyamoto, S., Ronsein, G.E., Medeiros, M.H.G., and Cadet, J. (2019). Singlet molecular oxygen reactions with nucleic acids, lipids, and proteins. Chem. Rev. *119*, 2043–2086.

72. Liu, F., Lu, W., Yin, X., and Liu, J. (2016). Mechanistic and kinetic study of singlet O_2 oxidation of methionine by on-line electrospray ionization mass spectrometry. J. Am. Soc. Mass Spectrom. *27*, 59–72.

73. Gracanin, M., Hawkins, C.L., Pattison, D.I., and Davies, M.J. (2009). Singlet-oxygen-mediated amino acid and protein oxidation: Formation of tryptophan peroxides and decomposition products. Free Radic. Biol. Med. *47*, 92–102.

74. Ehrenshaft, M., Silva, S.O., Perdivara, I., Bilski, P., Sik, R.H., Chignell, C.F., Tomer, K.B., and Mason, R.P. (2009). Immunological detection of *N*-formylkynurenine in oxidized proteins. Free Radic. Biol. Med. *46*, 1260–1266.

75. Pan, B., Abel, J., Ricci, M.S., Brems, D.N., Wang, D.I., and Trout, B.L. (2006). Comparative oxidation studies of methionine residues reflect a structural effect on chemical kinetics in rhG-CSF. Biochemistry *45*, 15430–15443.

76. Jensen, R.L., Arnbjerg, J., and Ogilby, P.R. (2012). Reaction of singlet oxygen with tryptophan in proteins: A pronounced effect of the local environment on the reaction rate. J. Am. Chem. Soc. *134*, 9820–9826.

77. Ghesquiere, B., Jonckheere, V., Colaert, N., Van Durme, J., Timmerman, E., Goethals, M., Schymkowitz, J., Rousseau, F., Vandekerckhove, J., and Gevaert, K. (2011). Redox proteomics of protein-bound methionine oxidation. Mol. Cell. Proteomics *10*, M110.006866.

78. Gray, H.B., and Winkler, J.R. (2015). Hole hopping through tyrosine/tryptophan chains protects proteins from oxidative damage. Proc. Natl. Acad. Sci. USA *112*, 10920–10925.

79. Levine, R.L., Mosoni, L., Berlett, B.S., and Stadtman, E.R. (1996). Methionine residues as endogenous antioxidants in proteins. Proc. Natl. Acad. Sci. USA *93*, 15036–15040.

80. Sreedhara, A., Lau, K., Li, C., Hosken, B., Macchi, F., Zhan, D., Shen, A., Steinmann, D., Schöneich, C., and Lentz, Y. (2013). Role of surface exposed tryptophan as substrate generators for the antibody catalyzed water oxidation pathway. Mol. Pharm. *10*, 278–288.

81. DePaz, R.A., Barnett, C.C., Dale, D.A., Carpenter, J.F., Gaertner, A.L., and Randolph, T.W. (2000). The excluding effects of sucrose on a protein chemical degradation pathway: methionine oxidation in subtilisin. Arch. Biochem. Biophys. *384*, 123–132.

82. Falconer, R.J., Chan, C., Hughes, K., Robert, J., Falconer, Cherrine, Chan, Karen, Hughes, and Munro, T.P. (2011). Stabilization of a monoclonal antibody during purification and formulation by addition of basic amino acid excipients. J. Chem. Technol. Biotechnol *86*, 942–948.

83. Arakawa, T., Tsumoto, K., Kita, Y., Chang, B., and Ejima, D. (2007). Biotechnology applications of amino acids in protein purification and formulations. Amino Acids *33*, 587–605.

84. Ha, E., Wang, W., and Wang, Y.J. (2002). Peroxide formation in polysorbate 80 and protein stability. J. Pharm. Sci. *91*, 2252–2264.

85. Liu, H., Gaza-Bulseco, G., and Zhou, L. (2009). Mass spectrometry analysis of photo-induced methionine oxidation of a recombinant human monoclonal antibody. J. Am. Soc. Mass Spectrom. *20*, 525–528.

86. Chumsae, C., Gaza-Bulseco, G., Sun, J., and Liu, H. (2007). Comparison of methionine oxidation in thermal stability and chemically stressed samples of a fully human monoclonal antibody. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. *850*, 285–294.

87. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

88. Kuhn, M. (2008). Building predictive models in R using the caret package. 28. 26, 10.18637/jss.v028.i05.

89. Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. *33*, 1–22.