European Journal of Pharmaceutical Sciences 97 (2017) 151–157

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Diversity selection, screening and quantitative structure–activity relationships of osmolyte-like additive effects on the thermal stability of a monoclonal antibody

Olubukayo-Opeyemi Oyetayo [a,*], Oscar Méndez-Lucio [b], Andreas Bender [b], Hans Kiefer [a]

[a] Institute of Applied Biotechnology, Biberach University of Applied Sciences, Biberach (Riss), Germany
[b] Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 28 July 2016
Received in revised form 11 November 2016
Accepted 14 November 2016
Available online 16 November 2016

Chemical compounds studied in this article:
Mannitol (PubChem CID: 6251)
Glutamine (PubChem CID: 5961)
Tricine (PubChem CID: 79,784)
L-Phenylalanine methyl ester hydrochloride (PubChem CID: 75,736)
Sarcosine ethyl ester hydrochloride (PubChem CID: 171,173)
1,4-Anhydroerythritol (PubChem CID: 641,773)

Keywords:
Osmolyte
Monoclonal antibody
Thermal stability
Protein stabilization
Differential scanning fluorimetry
Diversity selection
QSAR

ABSTRACT

Solvents used for therapeutic proteins in downstream processing and in formulations often contain stabilizing additives that inhibit denaturation and aggregation. Such additives are mostly selected based on their positive effect on thermal stability of the protein, and are often derived from naturally occuring osmolytes. To better understand the structural basis underlying the effect of additives, we selected a diverse library of compounds comprising 79 compounds of the polyol, amino acid and methylamine chemical classes and determined the effect of each compound on thermal stability of a monoclonal antibody as a function of compound concentration. Thermal stabilization of the antibody was influenced by solution pH. Quantitative structure–activity relationships (QSAR) were derived by partial least squares regression for individual compound classes and globally. The global model suggests that ligands with a phenyl ring will decrease the T_m, while highly soluble, polar compounds with at least two hydrogen bond donors will increase the T_m. This approach may be beneficial for further studies on the influence of other solution conditions like ionic strength and buffer species on additive-mediated protein stabilization.

1. Introduction

Osmolytes are small molecule compounds that stabilize the native structure of proteins. They occur naturally in organisms exposed to high temperature, extremes of ionic strength or pH, where intracellular proteins would otherwise unfold and aggregate (Macchi et al., 2012; Harries and Rösgen, 2008; Street et al., 2006). Some of these compounds are used as additives in the therapeutic protein production process and in formulations to stabilize the protein native state and to prevent aggregation (Macchi et al., 2012; Vagenende et al., 2009). Osmolytes were also used with specific aims of facilitating protein crystallization or preventing protein aggregation during production and in the life cycle of the finished product (Vedadi et al., 2006).

However, as osmolytes belong to different compound classes, the key molecular properties responsible for their protein-stabilizing effect are probably not uniform and overall only partially understood. In a situation, where only a small number of stabilizing additives identified by trial and error have made it into biopharmaceutical applications, the design of more potent stabilizers would strongly benefit from a better insight into the underlying principles of osmolyte-mediated protein stabilization, which is hence also the objective of the current study.

Abbreviations: mAb, Monoclonal antibody; MOE, Molecular Operating Environment; MDS, Multi-dimensional scaling; QSAR, Quantitative structure–activity relationship; DSC, Differential scanning calorimetry; DSF, Differential scanning fluorimetry; DMSO, Dimethyl sulfoxide; NaOH, Sodium hydroxide; HCl, Hydrochloric acid; LOO-CV, Leave-one-out cross-validation; PCA, Principal components analysis; PLS, Partial least squares; TMAO, Trimethylamine N-oxide; T_m, Melting temperature; TS potency, Thermal stabilization potency; VIP, Variable importance in projection; vs., Versus; RMSE, Root mean square error; Adj. R-squared, Adjusted R-squared.

* Corresponding author at: Hubertus Liebrecht 35, 88400 Biberach Riss, Baden-Wurttemberg, Germany.
E-mail address: oyetayophdhbc@gmail.com (O.-O. Oyetayo).

http://dx.doi.org/10.1016/j.ejps.2016.11.016
0928-0987/

Osmolytes can interact non-specifically or non-covalently with a solvated protein via hydrogen bonding, electrostatic interactions and indirectly by altering the water structure (Kamerzell et al., 2011). While more focus had been on non-specific interactions, protein-specific interactions with osmolytes have also been reported (Falconer et al., 2011). Different indirect measures of osmolyte–protein interactions include osmolality (Harries and Rösgen, 2008), the preferential interacting parameter (Arakawa and Timasheff, 1985), thermal conductivity (Park et al., 2011) and melting temperature (Kamerzell et al., 2011). However, a single osmolyte might interact with a solvated protein in multiple ways, and various biophysical methods must be applied to detect all of those interactions as each method has its inherent strengths and weaknesses (Kamerzell et al., 2011). Protein stabilization by osmolytes could then be achieved via various mechanisms such steric exclusion, cohesive force or surface tension effects, and the widely accepted preferential surface exclusion (Ohtake et al., 2011; Kumar et al., 2012)., However, it is not always clear to what extent these interactions or mechanisms contribute to the overall observed effect on the thermal stability of the protein.

Also, osmolyte effects had been linked to measured physical properties such as pKa (Falconer et al., 2011), viscosity (He et al., 2011a) and surface tension (Kaushik and Bhat, 1998). Many of these reports seek to identify general stabilization mechanisms of certain osmolyte classes irrespective of the protein sequence or source. The ubiquitous nature of the peptide backbone in all proteins is often exploited. However, there are mentions of protein-specific effects due to the presence of side groups (Falconer et al., 2011; Street et al., 2006; Harries and Rösgen, 2008). Computational and experimental studies aimed at rationalizing additive effects on protein stability either focus on a few selected additives like trehalose (Jain and Roy, 2009), proline (Ignatova and Gierasch, 2006), glycerol (Gekko and Timasheff, 1981) and TMAO (Ma et al., 2014) or on specific chemical moieties such as the guanidino group (Zarrine-Afsar et al., 2006). Other related studies are limited to either a class of osmolyte, e.g. polyols (Roughton et al., 2012; London et al., 1979), amino acids (Falconer et al., 2011; Taneja and Ahmad, 1994) methylamines (Arakawa and Timasheff, 1985) or to a limited number of the three classes (Macchi et al., 2012; Street et al., 2006).

In the present study, we present a systematic approach that focuses entirely on molecular properties of additives and does not postulate a specific mechanism of interaction. To that end, we selected a comprehensive compound library comprising amino acids, methylamines and polyols, and measured their effects on the thermal stability of a monoclonal antibody (mAb). The library was not restricted with respect to toxicity, compound stability or GMP compliance at the present stage, as we were primarily interested in identifying the molecular properties responsible for effects on protein stability. Such properties can be included in a later step, when the results will be used for the design of improved stabilizing additives. Measurement techniques for studying the thermal unfolding of proteins include circular dichroism, differential scanning calorimetry (DSC), nuclear magnetic resonance and differential scanning fluorimetry (DSF) using intrinsic protein fluorescence or extrinsic fluorescent probes (Kamerzell et al., 2011). We chose DSF as a measurement technique of thermal stability because of its high throughput capacity and the well-established correlation of DSF results with those of DSC (Ericsson et al., 2006; Niesen et al., 2007; Menzen and Friess, 2013). Drawbacks with respect to equilibrium methods such as DSC are subtle influences of the extrinsic fluorescent probe on T_m and the inability to measure the reversibility of unfolding transitions, resulting in apparent T_m values that do not necessarily reflect the equilibrium state. However, shifts of apparent T_m have been useful to detect changes in protein stability in many reports (He et al., 2010, 2011b). Since pH can affect protein conformation, experiments were carried out at two (three for polyols) pH values.

In a quantitative structure–activity relationship (QSAR) approach, the concentration-dependent effect of additives on the apparent T_m was correlated with physicochemical and structural molecular properties of compounds, coded as numerical descriptors. Roughton et al. described the utilization of QSARs in the design of carbohydrate excipients as aids for lyophilization (Roughton et al., 2012), thereby demonstrating the plausibility of this method. PLS regression yielded local and global models with satisfactory predictive statistics for interpolation. Variable importance in projection (VIP) (Chong and Jun, 2005) assessment of the global model revealed a strong dependence of observed effects on the polarity and charge of the compounds which are in accordance with most of previous findings, thereby demonstrating the capability of this approach. The results are encouraging for more extended studies including a variety of proteins, a broader formulation space and automated data requisition. This approach may also be explored to investigate osmolyte effects on colloidal stability since the conformational stability of proteins does not guarantee colloidal stability under the same conditions.

2. Experimental section

2.1. Materials

Disodium hydrogen phosphate (Na_2HPO_4), sodium chloride and citric acid monohydrate were purchased from Carl Roth GmbH (Karlsruhe, Germany). Sypro Orange at a concentration of 5000× in DMSO was purchased from Invitrogen GmbH (Darmstadt, Germany). Taurine, serine, proline and beta-alanine were purchased from Applichem (Darmstadt, Germany), other screening compounds were purchased from Sigma-Aldrich (Taufkirchen, Germany). All other reagents were of analytical grade.

2.2. Diversity selection of screening compounds

All amino acids with a molecular weight <300 $gmol^{-1}$ from the eMolecules database (James, 2011) were selected and sarcosine and mannitol were used as queries for the methylamine and polyol class, respectively. A molecular weight cutoff of <500 $gmol^{-1}$ and a Tanimoto similarity (Jaccard, 1912) of 0.5 was applied for the polyols. Molecular Access System (MACCS) keys were calculated with the Molecular Operating Environment (MOE) (ChemicalComputingGroupInc, 2004) and used to rank all the compounds in the master dataset based on molecular similarity and by Jarvis–Patrick clustering. Afterward, the most diverse compounds were selected; further pruning was done to exclude reactive and toxic compounds as indicated on the material safety data sheets. Compounds that were not soluble at or above 0.1 M in water were excluded; their solubility was assessed by visual observation. The final lists of the selected compounds can be found in the supporting information.

3. Methods

3.1. Buffer preparation

Citrate-phosphate buffers were prepared from 1 M citric acid monohydrate and 0.5 M disodium hydrogen phosphate based on a formula from the reference buffer table (Dawson et al., 1986). The buffer system was selected because of its constant buffering capacity, which spans from pH 2.6 to 7.6, and which allowed for the utilization of a single buffer system, thereby limiting the influence of differing buffer ions on the results obtained.

3.2. mAb preparation

A recombinant human monoclonal antibody of the IgG1 subclass (mAb1) was produced in-house in Chinese hamster ovary cells. The cell-producing line was obtained from Rentschler Biotechnologie GmbH (Laupheim, Germany). Purification was carried out with protein A affinity chromatography and subsequent sterile filtration. Before use,

mAb1 was dialyzed at 4 °C–8 °C overnight into a buffer of the desired pH with a buffer change after the first 4 h and filtered using 0.22 μm pore-sized filters.mAb1 concentrations were derived from absorbance measurements at 280 nm in Nanodrop 100 spectrophotometer (Peqlab Biotechnologie, Erlangen, Germany), with a calculated absorption coefficient of 1.47 ml mg^{-1} cm^{-1}. The theoretical isoelectric point of the mAb was 8.2.

3.3. Preparation of screening compound solutions

Thirty compounds were screened one day per week in a screening campaign that lasted eight weeks. Stock solutions of compounds with concentrations at their upper working concentrations in water were prepared in 5 ml Eppendorf tubes (Eppendorf, Wesseling-Berzdorf, Germany). To determine the upper working concentrations, an arbitrary cutoff of 0.5 M was set for all compounds. Poorly soluble compounds at 0.5 M were diluted further till complete dissolution and the approximate upper working concentration were determined. A total of 4 ml of filtered buffer stock was used to dissolve the compounds with vigorous mixing on a Vortex Genie® 2 mixer (Scientific Industries, New York, USA). The pH was then carefully adjusted with 2.5 M NaOH or 2.5 M HCl while mixing with a magnetic spin vanes (Sigma-Aldrich, Taufkirchen, Germany). pH checks were done with the use of a PerpHecT® Ross microelectrode (Thermo Scientific, Massachusetts, USA) coupled to an Orion VERSA STAR® pH meter (Thermo Scientific, Massachusetts, USA). The solutions were kept at 4 °C–8 °C and used up within three days of preparation in all cases. At the start of each screening day, 500 μl of each of the solutions were transferred into a 1.2 ml 96-well storage plate supplied with lids (Biozym Scientific, Oldendorf, Germany) in an order congruent to the reaction plate map.

3.4. Differential scanning fluorimetry (DSF)

The thermal stability of mAb1 was measured using DSF in a LightCycler 480 device (Roche, Mannheim, Germany) and with Sypro Orange as the fluorescent dye. Each plate included four negative controls without additive and four positive controls. The procedure was according to the protocol published by Niesen et al. (2007). Concentrations of compounds tested were 100%, 75%, 50% and 25% of the upper working concentration at pH 3.5 and pH 6.5. Thermal stabilization effects at pH 5.2 were also determined for the polyols. For pH conditions with more than one distinct melting transition, T_ms were read manually as the minima of inverse first derivative plots of the melting curve, generated using the LightCycler protein melting software version 1.4. Automatically reported values were used for conditions with only one distinct minimum.

The extracted T_m values, which correspond to the lowest observed thermal transition (T_m1), were converted to delta T_m values by subtracting the mean of 4 negative controls per plate from individual T_ms per well. Then, the derived delta T_m values for three replicates were averaged and the standard deviation calculated. The next step was to calculate the slopes of the plot of compound concentration versus delta T_m values. We termed those slopes “thermal stabilization potency (TS potency)” of the respective compounds. A linear fit was applied to all data, the intercepts were allowed to vary freely for T_ms at pH 3.5 but were fixed at zero for T_ms at pH 5.2 and 6.5.

3.5. Model derivation

Local 2D QSAR models were developed using 29 amino acids, 18 methylamines and 32 polyols as well as a global 2D QSAR model incorporating all 84 compounds. A set of 195 2D descriptors was calculated from the chemical structures of these compounds using the MOE software. Descriptors with low variance (<0.1) were removed, and the remainder were scaled and centered. This final set of descriptors was used as independent variables, whereas TS potency at pH 6.5 was the dependent variable. The models reported in this work were trained using partial least squares (PLS) as encoded in the PLS package (Mevik and Wehrens, 2007) in R statistical software. Leave-one-out cross-validation was used to evaluate the consistency (and to some extent predictive power) of the model. It is important to keep in mind that good performance during validations (in this case LOO-CV) is a necessary condition but is not always sufficient to ensure predictive power (Golbraikh and Tropsha, 2002). For model interpretation, all variables were ranked based on their variable importance in projection (VIP) as implemented by Chong and Jun (2005). All models reported are the results of the second iteration step after variables with low VIP (<1) from the first iteration had been excluded.

4. Results

4.1. Melting temperature derivation and behavior at different pH

The thermal stability of mAb1 was investigated by DSF as a function of pH. Basing on the initial Sypro Orange fluorescence at 20 °C, the antibody appears to be largely unfolded at pH 3 relative to other tested pH conditions (Fig. 1A). Inverse first derivative plots of the melting curves (Fig. 1B) allow the extraction of the apparent melting temperatures. At pH 3, the melting transitions did not yield distinct and sharp minima, but from pH 3.5 up to 4.8, two distinct minima were observed, whereas we could reckon with only one distinct minimum from pH 5 upwards. There appears to be another melting transition between 60 °C and 85 °C under most pH conditions tested, but this was disregarded and considered unreliable for a screening campaign. The midpoint temperature of the lowest observed thermal transition was used for QSAR

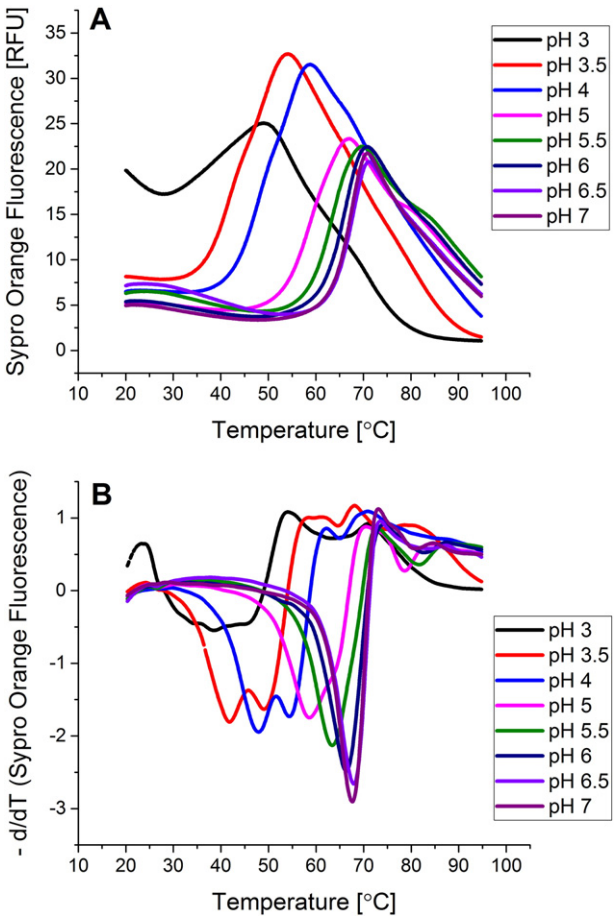

Fig. 1. (A) DSF melting transitions of mAb1 in citric-phosphate buffer at different pH values; (B) inverse first derivative plots of the thermal transitions for extraction of melting temperatures.

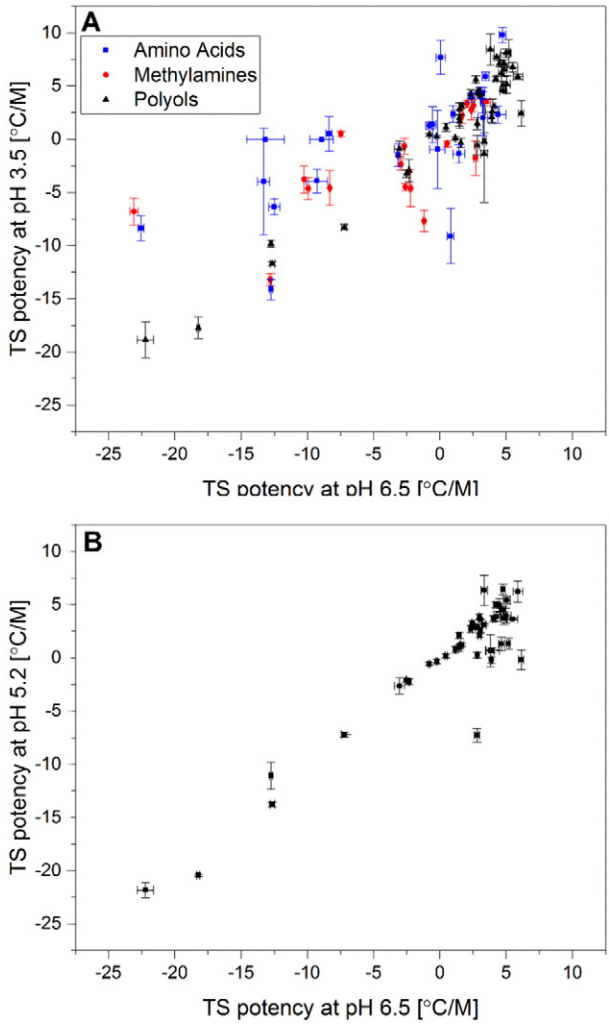

Fig. 2. Plots depicting (A) relationships between TS potencies of different classes of compounds at pH 6.5 and 3.5; (B) the relationship between TS potencies of polyols at pH 6.5 and 5.2 (error bars are the standard deviation calculated from three replicate measurements).

derivation in all cases. An increasing thermal stability of mAb1 with an increase in solution pH was observed.

4.2. TS potency of additives and effect of pH

Since pH clearly had an influence on the thermal stability of the mAb in base formulation buffer, the next step was to study the influence of pH on the TS potencies of different classes of compounds tested. pH 3.5 and pH 6.5 were selected as representatives of measurable extremes of folded and unfolded states of mAb1. It was found that while only ~25% of the methylamines and ~50% of the amino acids tested acted as stabilizers at either pH, this was the case for ~75% of the polyols (Fig. 2A). Both amino acids and methylamine classes showed poor correlation between TS potencies at the two different pH values. The polyol class seems to be largely unaffected in their TS potencies by solution pH (Fig. 2A), including pH 5.2 (Fig. 2B). The TS potencies at pH 6.5 had higher reproducibility than TS potencies at pH 3.5. The relative standard deviation of triplicate measurements for all compounds tested were

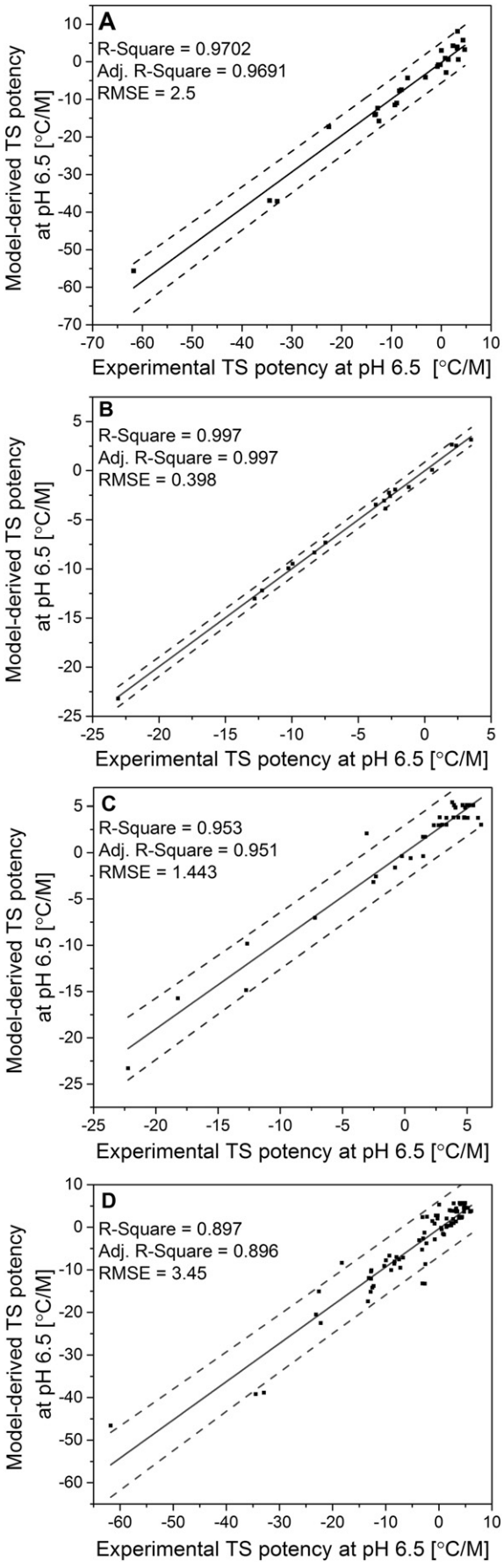

Fig. 3. Comparison of experimental TS potencies at pH 6.5 to model derived values during model training for (A) the amino acids, (B) the methylamines,(C) the polyol class and (D) the global PLS model comprising compounds in the amino acid, methylamine and polyol classes.

Table 1
Leave-one-out cross-validation results for PLS models. Independent models presented slightly better performance compared to the global model.

Model statistic	Amino acids	Methylamines	Polyols	Global
Sample size	29	18	37	84
Descriptors	63	56	51	60
R-squared value	0.864	0.97	0.906	0.679
Adj. R-squared value	0.848	0.967	0.891	0.645
RMSE	5.34	1.15	2.04	6.07
Components	9	10	3	4

lower than 10% and only 6 compounds had a relative standard deviation above 5%, indicating robustness of the assay performed.

4.3. QSAR model derivation and interpretation

Only TS potencies at pH 6.5 were used for QSAR analysis because linear fits through the origin gave better correlation coefficients than at pH 3.5, resulting in higher data quality. Also, pH 6.5 is closer to the pH used in most process steps, thus allowing direct interpretation of the derived model with respect to osmolyte function in biopharmaceutical applications.

QSAR models developed for each compound class as well as the global model (for all compounds) are shown in Fig. 3. Fits during model training (Fig. 3) and leave-one-out cross-validation (Table 1) showed high correlation values and were thus considered acceptable for interpolation. However, the model for amino acids was less accurate than the polyol and methylamine models, having a lower Adj. R-squared value and a higher RMSE. The global QSAR training model (Fig. 3D) had satisfactory performance during cross-validation, i.e. Adj. R-squared = 0.64 and $RMSE_{LOO}$ = 6.07 as shown in Table 1. In other words, it is well suited to predict the TS potencies of new compounds within the applicability domain. Table 2 enlists the ten most important variables for this model and their respective VIP. In general, the TS potencies at pH 6.5 increased when compounds present high values of logS, as well as with low values of SlogP (a particular implementation aimed at predicting logP partition coefficients) as shown in Fig. 4. In the same way, compounds with a higher fractional polar surface area (Q_VSA_FPOL) present higher TS potency, whereas compounds with high fractional hydrophobic surface area (Q_VSA_FHYD) act less stabilizing or even destabilizing (Fig. 4). Interestingly, the three compounds with the lowest TS potencies are those with a phenyl ring in the structure. These three compounds also present the highest values of atomic contribution to SlogP between 0.25 and 0.3 (SlogP_VSA7, 88.22), van der Waals surface are with a partial charge between −0.1 and −0.05 (PEOE_VSA-1, 61.27), number of aromatic atoms (a_aro, 6) and number of aromatic atoms (b_aro, 6). This could be related to the fact that all compounds with two or more hydrogen bond donors presented a delta $T_m > -10$. Based on the global model, we can suggest that a phenyl ring in the ligand will decrease the T_m; highly soluble, polar compounds with low SlogP and with at least two hydrogen bond donors will increase the T_m.

5. Discussion

5.1. pH effects on T_ms and TS potencies

pH is an important solution factor that determines the conformational state of a protein, since as the pH drops from near neutral conditions to acidic ones, the mAb gets increasingly charged, and it is expected that the mAb becomes readily susceptible to thermal denaturation due to electrostatic intramolecular repulsion forces. Accordingly, in earlier work reported from our lab (Oyetayo and Kiefer, 2016), mAb1 was shown to exist in increasingly unfolded states with decreasing pH based on Bis-ANS extrinsic fluorescence measurements. This is consistent with the observed Sypro Orange fluorescence at the start of each thermal transition (Fig. 2A), which we can also observe in the current study.

Since pH affects both protein charge and conformational stability, we also expect it to influence the thermal stabilization effects of additives. The latter will result both from changes in electrostatic interactions between charged compounds and charged amino acid side chains and from different selective compound interactions with the folded or the denatured state of the protein, respectively. Our results confirm pH dependency only for charged compounds, i.e. the amino acid and methylamine classes, where thermal stabilization effects at pH 3.5 and at pH 6.5 are poorly correlated (Fig. 2A). Polyols, on the other hand, seem to have largely similar effects on mAb1 thermal stability for all pH tested, i.e. pH 3.5, 5.2 and 6.5 (Fig. 2A and B). The results suggest that pH-dependent differences in stabilization potency can be attributed largely to electrostatic interactions, while the preference for folded vs. unfolded protein conformations appear to be independent of pH. This view confirms many earlier studies that came to the conclusion that protein stabilization by polyols and other osmolytes is often dominated by surface exclusion (Macchi et al., 2012; Kumar et al., 2012). This entropic effect will try to compress protein volume independent of its conformational state. Another general finding is the saturation of positive thermal shifts at high additive concentration, which was found under all conditions tested, independent of the additive class (Figs. 1B and 2). In their work, Cimmperman et al. addressed this effect in detail (Jachimovic et al., 2008).

5.2. Model interpretation

The 2D-QSAR models presented were based on 2D descriptors calculated from MOE. 2D molecular descriptors are numerical properties that can be calculated from the connection table representation of a molecule including elements, formal charges and bonds, but not atomic coordinates. These descriptors are, therefore, not dependent on the three-dimensional geometry of a molecule. They can be grouped as

Table 2
Variables with the highest VIP in the global PLS model. It can be seen that many of the variables with high VIP are associated with hydrophobicity or polarity.

Variable name	Description	VIP	Regression coefficient
SlogP_VSA7	Sum of the accessible surface area (in Å^2) over all atoms *i* such that SlogP of atom *i* is in (0.25, 0.30]	2.027	−1.127
PEOE_VSA-1	Sum of the accessible surface area (in Å^2) over all atoms *i* such that the partial charge of atom *i* is in [−0.10, −0.05)	1.929	0.678
logS	Log of aqueous solubility (mol/L)	1.798	−0.522
a_aro	Number of aromatic atoms	1.726	−0.438
b_ar	Number of aromatic bonds	1.726	−0.351
Q_VSA_FPOL	Fractional polar van der Waals surface area	1.709	−0.351
Q_VSA_FHYD	Fractional hydrophobic van der Waals surface area	1.709	−0.345
Q_VSA_HYD	Total hydrophobic van der Waals surface area	1.651	0.201
SMR_VSA5	Sum of the accessible surface area (in Å^2) overall atoms *i* such that the molar refractivity of atom *i* is in (0.44, 0.485]	1.632	−0.201
SlogP	Log of the octanol/water partition coefficient as an atomic contribution model	1.599	0.183

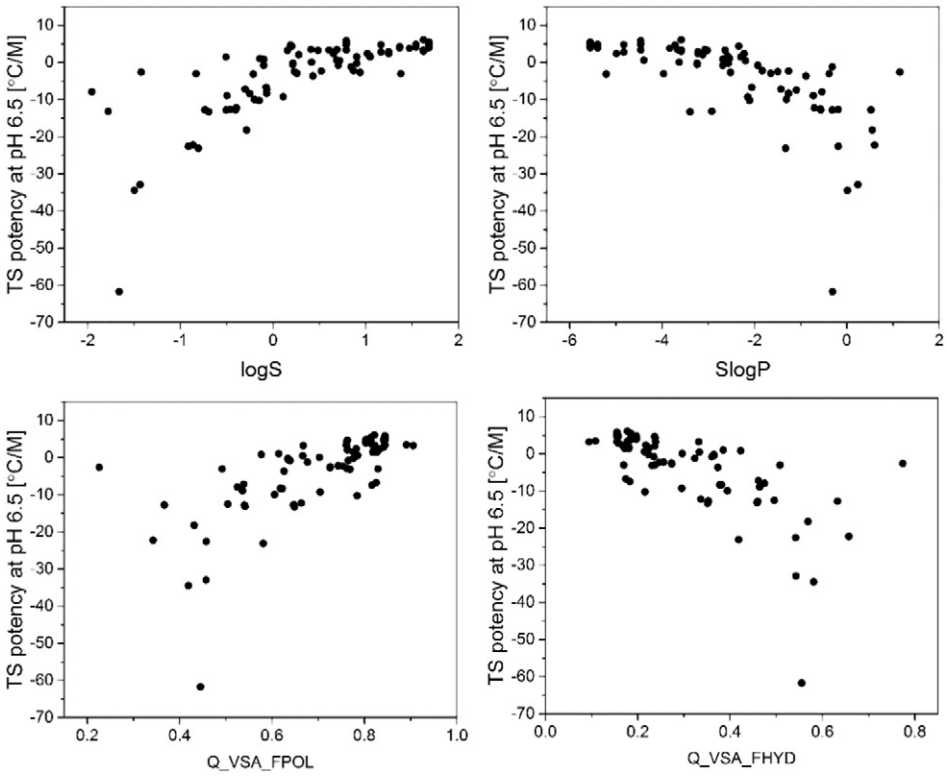

Fig. 4. Correlation of observed TS potencies at pH 6.5 with logS, SlogP, Q_VSA_FPOL and Q_VSA_FHYD for all compounds tested.

descriptors of physical properties, sub-divided surface areas, atom counts and bond counts, Kier and Hall connectivity and Kappa shape indices, adjacency and distance matrix descriptors, pharmacophore feature descriptors and partial charge (ChemicalComputingGroupInc, 2004). While we could derive QSARs by various methods and with other descriptors, the derived local models (as well as the global model) are of satisfactory performance as well as being internally consistent, as confirmed by a leave-one-out cross-validation (Table 1). This means that in the given variable space employed, comprising the physicochemical properties above, relationships between molecular properties and stabilization effects do exist and can be modelled.

VIP evaluation of the model for the most important variables (Table 2) hint at polarity or hydrophobicity as an important property for osmolyte effects on protein stability. Significant correlations between key descriptors and TS potency at pH 6.5 are shown in Fig. 4. A study by Taneja and Ahmad (1994) presented a hypothesis that protein stabilization by amino acids depends on the fine balance between the destabilizing preferential interaction of hydrophobic amino acids with unfolded protein and the stabilizing preferential exclusion of hydrophobic amino acids from the folded protein. They argued that upon denaturation, the buried hydrophobic side chains of the protein are exposed thereby favoring destabilizing preferential interactions with increasing amino acid hydrophobicity (Taneja and Ahmad, 1994). In a more detailed study of polyol-induced protein stability by molecular dynamics simulations (Liu et al., 2010), protein stability was shown to increase with an increase in polyol molecular volume and fractional polar surface area. This effect was explained to be based on the preferential exclusion theory and more influenced by the polyol size, with bigger polyols being more excluded from the protein surface. They attributed little effect to direct hydrogen bonds between the protein and the polyol (Liu et al., 2010).Street et al. (2006) investigated the idea of a universal underlying mechanism or property for stabilization across all osmolyte classes, howbeit with a limited number of compounds and with generalizations based only on the polar protein backbone. Contrary to our observation, they observed a negative trend in osmolyte stabilization effect with increasing fractional polar surface area. Provided that there are conflicting reports on the relationship between polarity and protein stability, it may be tangible to speculate that osmolytes impact on protein stability in one way or another via direct or indirect hydrogen-bonding interactions due to the ubiquitous nature of hydrogen bonds within solvents, proteins and most osmolytes (Kamerzell et al., 2011). Some of the other important variables (Table 2) refer to the accessible surface area of the osmolyte bearing a defined charge or polarity. For example, compounds having a neutral accessible surface area (charge between −0.1 and −0.05; high values of PEOE_VSA-1) are more likely to present low values of TS potency. In a similar way, compounds having a hydrophobic accessible surface area with low logP values (high values of SlogP_VSA7) are associated with low TS potency values. While the reported QSAR approach does not provide detailed mechanistic insight into the mechanism of stabilization by osmolytes, it could serve as a tool in directing further research on osmolyte's effects on protein stability. The derived models can now be used for the rapid selection of untested additives in order to design improved additives in an iterative manner, which will also be our objective for future work.

Acknowledgment

The Federal Ministry of Education and Research, Germany (BMBF) funded this work (grant number 0315338A). Thanks are due to Natalie Nilson and Julia Czwikla for assistance during the screening campaign. The authors acknowledge the support and inputs of Bernd Burghardt, Maike Eisenkolb, Fabian Bickel, Qian Zhou and Vivek Thakare at the preliminary and exploratory phases of the project.

Appendix A. Supplementary data

The lists of compounds screened and used for QSAR model training in Excel format. Supplementary data associated with this article can be found in the online version, at doi: http://dx.doi.org/10.1016/j.ejps.2016.11.016.

References

Arakawa, T., Timasheff, S.N., 1985. The stabilization of proteins by osmolytes. Biophys. J. 47, 411–414.

ChemicalComputingGroupInc, 2004. Molecular Operating Environment (MOE). Sci. Comput. Instrum. 22, 32.

Chong, I.G., Jun, C.H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemom. Intell. Lab. Syst. 78, 103–112.

Dawson, R.M.C., Elliot, D.C., Elliot, W.H., 1986. In: Jones, K.M. (Ed.), Data for Biochemical Research. Oxford Sci. Publ. (at <http://www.sigmaaldrich.com/life-science/core-bioreagents/biological-buffers/learning-center/buffer-reference-center.html#citric>).

Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., Nordlund, P., 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal. Biochem. 357, 289–298.

Falconer, R.J., Chan, C., Hughes, K., 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.

Gekko, K., Timasheff, S.N., 1981. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol–water mixtures. Biochemistry 20, 4667–4676.

Golbraikh, A., Tropsha, A., 2002. Beware of q2! J. Mol. Graph. Model. 20, 269–276.

Harries, D., Rösgen, J., 2008. A practical guide on how osmolytes modulate macromolecular properties. Methods Cell Biol. 84, 679–735.

He, F., Hogan, S., Latypov, R.F., Narhi, L.O., Razinkov, V.I., 2010. High Throughput Thermostability Screening of Monoclonal Antibody Formulations. 99 pp. 1707–1720.

He, F., et al., 2011a. Screening of Monoclonal Antibody Formulations Based on High-Throughput Thermostability and Viscosity Measurements: Design of Experiment and Statistical Analysis. 100 pp. 1330–1340.

He, F., et al., 2011b. Screening of monoclonal antibody formulations based on high-throughput thermostability and viscosity measurements: design of experiment and statistical analysis. J. Pharm. Sci. 100, 1330–1340.

Ignatova, Z., Gierasch, L.M., 2006. Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant \r10.1073/pnas.0603772103. PNAS 103, 13357–13361.

Jaccard, P., 1912. The distribution of the flora in the alphine zone. New Phytol. XI, 37–50.

Jachimovic, S., Bumelis, V., Matulis, D., 2008. A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. 95, 3222–3231.

Jain, N.K., Roy, I., 2009. Effect of trehalose on protein structure. Protein Sci. 18, 24–36.

James, C.A., 2011. Chemoinformatics 101. eMolecules.

Kamerzell, T.J., Esfandiary, R., Joshi, S.B., Middaugh, C.R., Volkin, D.B., 2011. Protein–excipient interactions: mechanisms and biophysical characterization applied to protein formulation development. Adv. Drug Deliv. Rev. 63, 1118–1159.

Kaushik, J.K., Bhat, R., 1998. Thermal stability of proteins in aqueous polyol solutions: role of the surface tension of water in the stabilizing effect of polyols. J. Phys. Chem. B5647, 7058–7066.

Kumar, V., Singh, S.N., Kalonia, D.S., 2012. Mechanism of stabilization of proteins by polyhydroxy co-solvents: concepts and implications in formulation development. Am. Pharm. Rev. 15.

Liu, F.-F., Ji, L., Zhang, L., Dong, X.-Y., Sun, Y., 2010. Molecular basis for polyol-induced protein stability revealed by molecular dynamics simulations. J. Chem. Phys. **132**, 225103.

London, W.N., Arch, K., Back, J.F., Oakenfull, D., Smith, M.B., 1979. Increased Thermal Stability of Proteins in the Presence of Sugars and Polyolst. 18 pp. 5191–5196.

Ma, J., Pazos, I.M., Gai, F., 2014. Microscopic Insights into the Protein-Stabilizing Effect of Trimethylamine N-Oxide (TMAO). http://dx.doi.org/10.1073/pnas.1403224111.

Macchi, F., Eisenkolb, M., Kiefer, H. & Otzen, D. E. The effect of osmolytes on protein fibrillation. 3801–3819 (2012). doi:10.3390/ijms13033801

Menzen, T., Friess, W., 2013. High-throughput melting-temperature analysis of a monoclonal antibody by differential scanning fluorimetry in the presence of surfactants. J. Pharm. Sci. 102, 415–428.

Mevik, B.H., Wehrens, R., 2007. The PLS package: principal component and partial least squares regression in R. J. Stat. Softw. 18, 1–23.

Niesen, F.H., Berglund, H., Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat. Protoc. 2, 2212–2221.

Ohtake, S., Kita, Y., Arakawa, T., 2011. Interactions of formulation excipients with proteins in solution and in the dried state. Adv. Drug Deliv. Rev. 63, 1053–1073.

Oyetayo, O.-O., Kiefer, H., 2016. Experimental Model System to Study pH Shift-Induced Aggregation of Monoclonal Antibodies under Controlled Conditions. Pharm. Res, Advanve On.

Park, B.K., Park, J., Lee, J.Y., 2011. Thermal Conductivity of Bovine Serum Albumin: A Tool to Probe Denaturation of Protein.

Roughton, B.C., Topp, E.M., Camarda, K.V., 2012. Use of glass transitions in carbohydrate excipient design for lyophilized protein formulations. Comput. Chem. Eng. 36, 208–216.

Street, T.O., Bolen, D.W., Rose, G.D., 2006. A molecular mechanism for osmolyte-induced protein stability. Proc. Natl. Acad. Sci. U. S. A. 103, 13997–14002.

Taneja, S., Ahmad, F., 1994. Increased thermal stability of proteins in the presence of amino acids. Biochem. J. 303 (Pt 1), 147–153.

Vagenende, V., Yap, M.G.S., Trout, B.L., 2009. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry 48, 11084–11096.

Vedadi, M., et al., 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proc. Natl. Acad. Sci. U. S. A. 103, 15835–15840.

Zarrine-Afsar, A., Mittermaier, A., Kay, L.E., Davidson, A.R., 2006. Protein stabilization by specific binding of guanidinium to a functional arginine-binding surface on an SH3 domain. Protein Sci. 15, 162–170.