Proton affinity of several basic non-standard amino acids

Marko Rožman

A B S T R A C T

The structures and absolute proton affinities of several arginine (2-amino-3-guanidinopropionic acid, 2-amino-4-guanidinobutyric acid, homoarginine, citrulline and canavanine), histidine (1-methylhistidine and 3-methylhistidine) and lysine (2,3-diaminopropanoic acid, 2,4-diaminobutanoic acid, ornithine, 5-hydroxylysine, canaline and thialysine) homologues and analogues have been estimated using compos-ite G3MP2B3 computational protocol. For a majority of here studied non-standard amino acids the gas-phase proton affinities were established for the first time, while for the others obtained values are used to improve the accuracy of the computational and experimental proton affinities reported previ-ously. In addition, structures and proton affinities are discussed in order to rationalize their biological activity.

1. Introduction

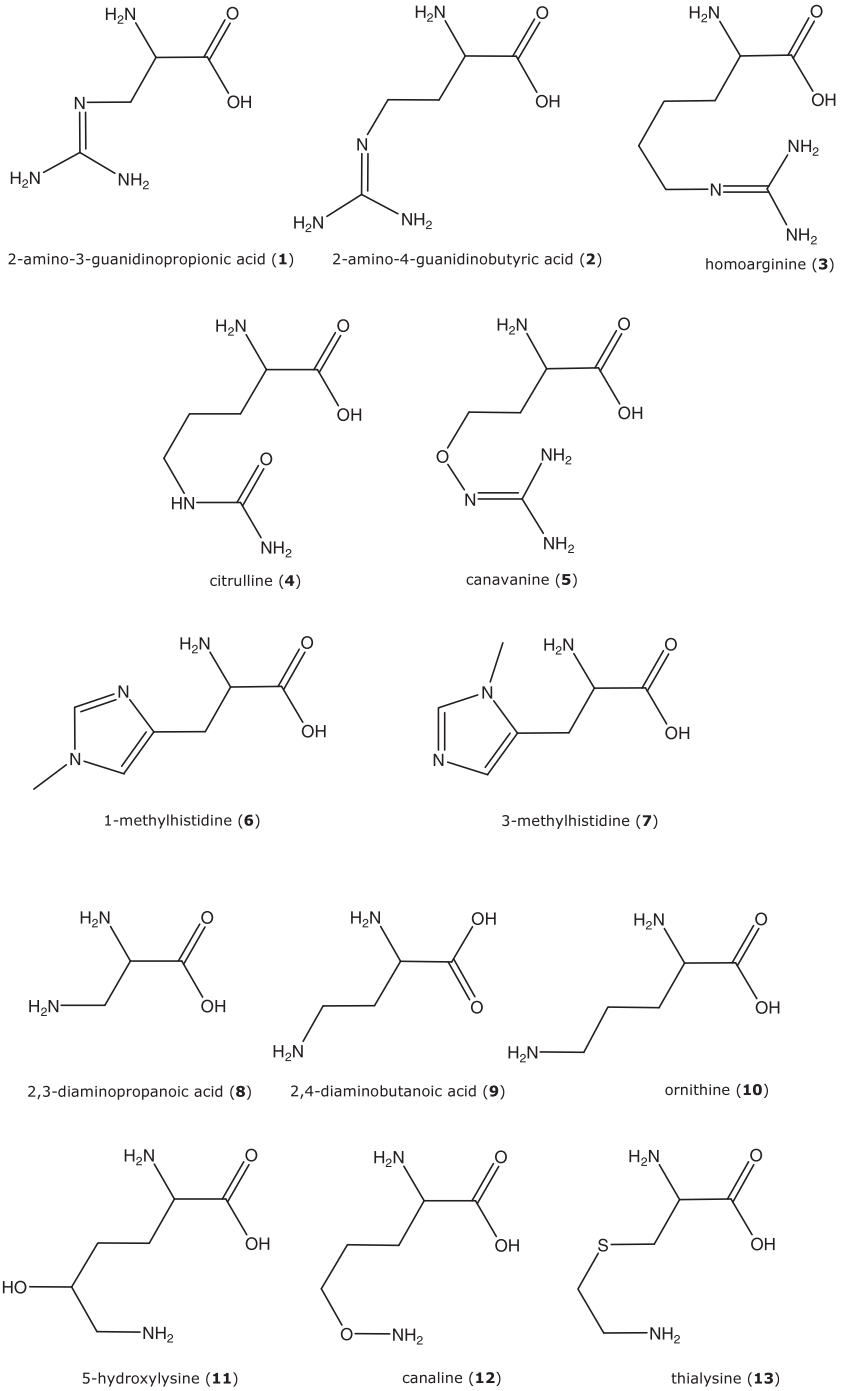
Amino acids, small biomolecules, are the principal building blocks of proteins and other biomolecules. Basically all organisms construct their proteins from only twenty amino acids, encoded by the genetic code. After biosynthesis, many proteins experience post-translational modifications which profoundly change the chemical nature of an amino acid (e.g. phosphorylation) and thus affect their structure, function, and activity. Post-translation modifications of proteinogenic amino acids is just one way of becoming a member of the much larger family called non-standard (natural and synthetic) amino acids. The ability to change nature of amino acid or to incorporate a non-standard amino acid (NSAA) into peptide and (or) protein enables creation of novel molecules with interesting chemical properties [1,2]. NSAA are associated with new enzymes, protein based therapeutics as well as with improvements in mass spectrometric characterization of peptides/proteins [1–4].

In order to create protein with new chemical properties or change the peptide bond dissociation process one should be familiar with a variety of intrinsic properties of NSAA such as gas-phase structure, heats of formation and energies associated to ionization processes e.g. proton affinity (PA). Intrinsic structural and thermochemical properties of standard amino acids have been extensively investigated and re-evaluated [5–22] (recent overview of the gas-phase thermochemistry of standard amino acids can be found in Ref. [22]). However, this is not the case with properties of NSAA. Poutsma et al. reported PA and the gas-phase acidity of several arginine, lysine and proline homologues determined experimentally by using the extended kinetic method [23–26]. Reported experimental values have been supported by density functional theory calculations at the B3LYP/6-311++G**//B3LYP/6-31+G* level.

The present Letter aims to extend conformational and thermochemical data of NSAA by considering a larger group of basic NSAA consisting of arginine, histidine and lysine analogues and homologues, Scheme 1. Detailed scans of the conformation surfaces of the protonated and neutral forms by means of composite G3MP2B3 method are presented. PA of studied NSAA are provided in order to: (i) improve on the accuracy of the computational and experimental PA reported by Poutsma and co-workers (canavanine (**5**), 2,3-diaminopropanoic acid (**8**), 2,4-diaminobutanoic acid (**9**), ornithine (**10**) and canaline (**12**) in Scheme 1), (ii) establish for the first time accurate gas-phase PA of 2-amino-3-guanidinopropionic acid (**1**), 2-amino-4-guanidinobutyric acid (**2**), homoarginine (**3**), citrulline (**4**), 1-methylhistidine (**6**), 3-methylhistidine (**7**), 5-hydroxylysine (**11**) and thialysine (**13**), Scheme 1.

2. Computational methods

The potential energy surface of the respective protonated and neutral amino acids was sampled by the conformational search engine implemented in HyperChem (Waterloo, Ontario, Canada). The method involves random variation of selected dihedral angles to generate new structure followed by energy minimization with the AM1 method. Each initial structure was selected by the usage directed scheme where all previously accepted conformations are sampled [27]. Unique low-energy conformations are stored while high-energy or duplicate structures are discarded. One thousand structures were generated with this procedure for every amino acid variant. Structures within $\sim$40 kJ mol^{-1} from the lowest energy conformer were selected, manually inspected and grouped



Scheme 1. Studied non-standard amino acids

into families. Up to 20 lowest conformations per each amino acid form were used as starting points for optimization at the B3LYP/6-31G(d) level of theory. The B3LYP/6-31G(d) total electronic energies were obtained and low-energy conformations (energy cut-off of 15 kJ mol^{-1} relative to the lowest was applied) were then submitted to G3MP2B3 composite method [28]. Finally, conformers were ranked according to the obtained G3MP2B3 Gibbs free energy (G°_{298}). No effort was made to improve the quality of Gibss free energy calculations in Gaussian.

The most stable conformer PA (PA_{SC}) and averaged PA (PA_{AV}) values are reported. PA_{SC} was calculated considering the most stable (neutral and protonated) conformer at 298 K while PA_{AV} was calculated considering the ensemble of conformers at 298 K for both neutral and protonated NSAA. The molar fractions of each conformer were calculated assuming a Boltzmann distribution (which should be only considered approximate, due to the use of the harmonic approximation), more details about protocol can be found in references [11,22]. All quantum mechanics data have been obtained by using the Gaussian 03 and 09 suite programs [29].

3. Results and discussion

3.1. Arginine homologues and oxy-analogues

The lowest energy conformers of neutral arginine homologues and oxy-analogues are dominated by strong hydrogen bonding between carboxylic hydrogen and guanidino nitrogen (ureido group

Table 1
Computed and experimental proton affinities (kJ mol^{-1}) of arginine homologues and analogues.

NSAA	PA_{SC}	PA_{AV}	Poutsma – computed[a]	Δ[b]	Poutsma – experimental[a]	Δ[b]
2-Amino-3-guanidinopropionic acid	1025.9	1026.5				
2-Amino-4-guanidinobutyric acid	1040.4	1035.9				
Homoarginine	1047.1	1046.1				
Citrulline	985.3	983.3				
Canavanine	1005.2	1009.5	1014	−8.8 (−4.5)	1001 ± 9	+4.2 (+8.5)

[a] Values taken from Ref. [25].
[b] Deviation relative to PA_{SC} (PA_{AV}) G3MP2B3 computed value.

oxygen in citrulline (**4**)); Figures 1, 3, 5, 7, 9 in Supporting material. This strong internal stabilization causes cyclic arrangements in all lowest energy neutral forms. Protonated species also have cyclic structures, however, with multiple strong hydrogen bonds involving guanidino, amino and carboxyl groups, (Figures 2, 4, 6, 8, 10 in Supporting material). In protonated structures the proton is placed on the guanidino group. However, in citrulline (**4**) due to the modification of the guanidino to the ureido group, the α-amino group becomes more basic and thus protonation site.

Determined PA values of arginine homologues suggest an increase in PA with the side chain length, Table 1. Observed effect may be related to a formation of a more efficient internal hydrogen bonding, particularly for the protonated forms, due to increase in the side chain length. More efficient intramolecular hydrogen bonding should reflect as larger negative entropy of protonation. Similar observations can be found for α,ω-diamines and lysine homologues [9,24]. Increase in PA diminishes at homoarginine (**3**) suggesting saturation effect when the guanidino group is at ε-position, similar as for lysine homologues (*vide infra*). High PA of homoarginine (**3**) corresponds with experimental findings of the increased basicity of lysine residues by conversion to homoarginine (**3**) [3]. High PA of homoarginine (**3**) converts the properties of lysine-containing peptides in tandem mass spectrometry experiments and thus increases database search selectivity [3].

Citrulline (**4**) and canavanine (**5**), oxy analogues of arginine exhibit decrease in PA due to modification of guanidino group (citrulline (**4**)) and inductive effect of the oxygen atom (canavanine (**5**)). Modification of the guanidino to the ureido group causes decrease in basic character of the side chain and affects the PA by ∼61 kJ mol^{-1} (PA of arginine 1046.4 kJ mol^{-1}, from Ref. [20]). Decrease in PA leads to uncharged side chain and has important consequences for the structure and function of proteins due to changes in protein folding [30]. In canavanine (**5**) substituted oxygen atom due to its electron withdrawing nature destabilizes positive charge on the side chain guanidino group, resulting in ∼41 kJ mol^{-1} decrease in PA, comparing to arginine. This finding is consistent with results showing that structurally aberrant proteins (containing canavanine instead of arginine) do not function properly [31]. Obtained PA_{SC} of canavanine (**5**) is in good agreement with Poutsma's experimental value, (difference of 4.2 kJ mol^{-1}, Table 1) [25]. Poutsma et al. reported larger discrepancy (13 kJ mol^{-1}) between theoretical and experimental PA and suggested that experimental PA of canavanine (**5**) could be underestimated [25]. However, we believe that his experimental value is more reliable especially due to the fact that a new, energetically more favored neutral structure was identified (Figure 9, Supporting material).

3.2. N-methylhistidines

Histidine has the third highest proton affinity among proteinogenic amino acids. Calculated proton affinity of histidine varies between 970 and 978.6 kJ mol^{-1} [17,19,20] while experimental values are in the range from 959 to 995.8 kJ mol^{-1} [8,9,12–15]. Pronounced basic character of histidine is due to the imidazole side chain. In order to modulate proton affinity and intramolecular hydrogen bonding between the amino group and the imidazole ring nitrogen sites two structural isomers, 1-methylhistidine (**6**) and 3-methylhistidine (**7**) were created. The idea was to modulate hydrogen bonds and consequently proton affinity by creating a *N*-methyl group which will serve as pi-electron donor and on the other hand represent a steric hindrance to formation of hydrogen bonds. PA_{SC} values of 993.9 (PA_{AV} = 993.1 kJ mol^{-1}) and 983.4 kJ mol^{-1} (PA_{AV} = 984.9 kJ mol^{-1}) estimated for 1-methylhistidine (**6**) and 3-methylhistidine (**7**), respectively, are higher than of hisitdine (975.7 kJ mol^{-1} from Ref. [16]). Both systems benefit from *N*-methylation which donates to the pi-electron system and provides additional stabilization for positive charge on the other nitrogen. However, methylation of N-3 poses a steric hindrance on the more stable protonation site (N-3) which is unable to form a strong intramolecular hydrogen bond with the amino group (Figure 14, supporting information), known to be responsible for the stability of the protonated histidine [17]. Lack of additional stabilization reflects as higher energy on the conformation of protonated form and corresponds to a lower proton affinity of 3-methylhistidine (**7**), as compared to 1-methylhistidine (**6**).

The knowledge that histidine is a member of the catalytic triad (Asp-His-Ser), which is found inside the active sites of some protease enzymes, offers an opportunity to modulate enzyme activity. Thus, by introducing *N*-methylated histidine with higher proton affinity and(or) steric hindrance, the potential energy profile of an enzymatic reaction may be altered, as shown for pancreatic phospholipase [32].

3.3. Lysine homologues and hydroxy/oxy/thio analogues

PA of lysine homologues and hydroxy/oxy/thio analogues are reported in Table 2 along with experimental and theoretical data from Poutsma and co-workers [24]. The 'most stable conformer' PA average deviation from Poutsma theoretical values is 8.1 kJ mol^{-1} but only 3.2 kJ mol^{-1} from experimental ones, Table 2. Larger discrepancies between theoretical values can be attributed to accuracy of theoretical method (B3LYP vs. G3MP2B3) as well as to the use of different conformations. For example, PA_{SC} of 2,3-diaminopropanoic acid (**8**) calculated from identical structures differed by 7.8 kJ mol^{-1}. Besides using more accurate theoretical method, usually one more stable conformation (neutral or protonated) was identified; e.g. for protonated ornithine (**10**) Poutsma most stable conformation was 17 kJ mol^{-1} higher in energy (obtained with G3MP2B3 method). Overall, it seams that more accurate theoretical method and better conformational choices resulted with smaller deviation from experimental values.

The lowest energy conformers of neutral species are characterized by two types of structures: either with or without the side chain amino group interactions with a carboxylic or α-amino group. Ornithine (**10**), hydroxylysine (**11**) and canaline (**12**) have extended structures with no side chain interactions (Figures 19, 21, 23 supporting information). On the other hand, cyclic structure

Table 2
Computed and experimental proton affinities (kJ mol^{-1}) of lysine homologues and analogues.

NSAA	PA_{SC}	PA_{AV}	Poutsma – computed[a]	Δ[b]	Poutsma – experimental[a]	Δ[b]
2,3-Diaminopropanoic acid	948.1	947.7	940.3	+7.8 (+7.4)	950.2 ± 7.2	−2.1 (−2.5)
2,4-Diaminobutanoic acid	969.7	968	972.5	−2.8 (−4.5)	975.8 ± 7.4	−6.1 (−7.8)
Ornithine	1005.2	1005.8	993.8	+11.4 (+12)	1001.1.2 ± 6.6	+4.1 (+4.7)
5-Hydroxylysine	992.7	992.7				
Canaline	949.7	951.1	960.0	−10.3 (−8.9)	950.0 ± 7	−0.3 (+1.1)
Thialysine	995.3	999.6				

[a] Values taken from references [24,25].
[b] Deviation relative to PA_{SC} (PA_{AV}) G3MP2B3 computed value.

and side chain internal hydrogen bonding is present in 2,3-diaminopropanoic acid (**8**), 2,4-diaminobutanoic acid (**9**) as well as in thialysine (**13**) (Figures 15, 17, 25 supporting information). Interaction between two amino groups is present in 2,3-diaminopropanoic acid (**8**) and 2,4-diaminobutanoic acid (**9**) while in thialysine (**8**) the side chain amino group interact with the carboxylic group hydrogen. Thialysine (**13**) has the strongest interaction; indicated by a hydrogen bond of 1.8 Å as compared to 2.5 and 3 Å in 2,3-diaminopropanoic acid (**8**) and 2,4-diaminobutanoic acid (**9**), respectively.

Upon protonation, all conformations of protonated lysine homologues and hydroxy/oxy/thio analogues are characterized by stabilization of the cation via intramolecular hydrogen bonding. This fact suggests large negative cyclization entropy for ornithine (**10**), hydroxylysine (**11**) and canaline (**12**) due to extended structures found for their neutral forms. In 2,3-diaminopropanoic acid (**8**), 2,4-diaminobutanoic acid (**9**), ornithine (**10**), hydroxylysine (**11**) and thialysine (**13**) protonation occurs on the side chain and the charged amino group is stabilized by strong intramolecular hydrogen bonds. The same intramolecular arrangement with extensive charge solvation is found in lysine [18]. On the other hand, canaline (**12**) has its proton more tightly bound to the α-amino group, suggesting that the α-amino group is the more basic site (Figure 24, supporting information) The α-amino group is stabilized by two intramolecular hydrogen bonds, one from side chain amino group and the other from carbonyl group oxygen. Lower basic character of the side chain amino group can be explained by the electron withdrawing nature of the substituted oxygen atom. Inductive effect of the oxygen atom destabilizes positive charge on the side chain amino group, contributing to significantly lower PA of canaline (**12**) (55 kJ mol^{-1} lower than ornitine (**10**)). Much subtle inductive effects exhibit hydroxylysine (**11**) and thialysine (**13**), both having electron withdrawing group at three bonds distance from the side chain amino group. Hydroxylic group and sulphur atom substitution in the side chain decreased PA by 8 and 5 kJ mol^{-1} with respect to lysine (1000.4 kJ mol^{-1} from Ref. [19]). This is a nice example how inductive effects weaken as the distance from the substituent increases. Their slightly lower PA (with respect to lysine) and their ability to retain a proton on their side chain allows hydroxylysine (**11**) and thialysine (**13**) to retain ionic interaction and act either as stabilizer for intra- and intermolecular crosslinks (hydroxylysine (**11**)) [33] or as site-specific mutagen in phosphorylation mapping (thialysine (**13**)) [4].

PA values of lysine homologues (2,3-diaminopropanoic acid (**8**), 2,4-diaminobutanoic acid (**9**) and ornithine (**10**)) suggest relationship between structure and PA. Increase in PA is related to the side chain length, although a saturation effect is observed as PA value of lysine is slightly lower than PA of ornithine (**10**). Similar trend can be followed for α,ω-diamines, with two longer homologues having the same PA and two shorter having lower [9], as already noted by Poutsma and co-workers [24].

4. Conclusion

The letter presented here uses extensive conformational search followed by G3MP2B3 composite method to estimate proton affinities of several arginine, histidine and lysine homologues and analogues. Majority of the gas-phase proton affinities here were established for the first time. Obtained 'most stable conformer' proton affinity values (in kJ mol^{-1}) are: 1025.9 (2-amino-3-guanidinopropionic acid), 1040.4 (2-amino-4-guanidinobutyric acid), 1047.1 (homoarginine), 985.3 (citrulline), 1005.2 (canavanine), 237.5 (1-methylhistidine), 235 (3-methylhistidine), 948.1 (2,3-diaminopropanoic acid), 969.7 (2,4-diaminobutanoic acid), 1005.2 (ornithine), 992.7 (5-hydroxylysine), 949.7 (canaline) and 995.3 (thialysine). Excellent agreement was found with available experimental thermochemistry data; mean absolute deviation was 3.4 kJ mol^{-1}. The information on their conformational properties and PA was directly related to the modes of their biomolecular performance and enhancing capabilities in proteome analysis.

Acknowledgement

Helpful discussions with Laboratory members are gratefully acknowledged. The Ministry of Science, Education and Sports of Republic of Croatia supported this work (Grant No.: 098-0982915-2945).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cplett.2012.06.048.

References

[1] L. Wang, A. Brock, B. Herberich, P.G. Schultz, Science 292 (2001) 498.
[2] J.C. Jackson, S.P. Duffy, K.R. Hess, R.A. Mehl, J. Am. Chem. Soc. 128 (2006) 11124.
[3] A.G. Sullivan et al., Int. J. Mass Spectrom. 210–211 (2001) 665.
[4] Z.A. Knight, B. Schilling, R.H. Row, D.M. Kenski, B.W. Gibson, K.M. Shokat, Nat. Biotechnol. 21 (2003) 1047.
[5] G. Bojesen, J. Am. Chem. Soc. 109 (1987) 5557.
[6] G.S. Gorman, J.P. Spier, C.A. Turner, I.J. Amster, J. Am. Chem. Soc. 114 (1992) 3986.
[7] S. Bouchonnet, Y. Hoppilliard, Org. Mass Spectrom. 27 (1992) 71.
[8] A.G. Harrison, Mass Spectrom. Rev. 16 (1997) 201.
[9] E.P. Hunter, S.G. Lias, J. Phys. Chem. Ref. Data 27 (1998) 3.
[10] Z.B. Maksić, B. Kovačević, Chem. Phys. Lett. 307 (1999) 497.
[11] W. Sun, G.R. Kinsel, D.S. Marynick, J. Phys. Chem. A 103 (1999) 4113.
[12] C. Afonso, F. Modeste, P. Breton, F. Fournier, J.C. Tabet, Eur. J. Mass Spectrom. 6 (2000) 443.
[13] G. Bouchoux, J.Y. Salpin, Eur. J. Mass Spectrom. 9 (2003) 391.
[14] J.F. Gal, M. Stone, C.B. Lebrilla, Int. J. Mass Spectrom. 222 (2003) 259.
[15] G. Bouchoux, D. Buisson, C. Colas, M. Sablier, Eur. J. Mass Spectrom. 10 (2004) 977.
[16] M. Rožman, J. Mass Spectrom. 40 (2005) 1357.
[17] B. Kovačević, M. Rožman, L. Klasinc, D. Srzić, Z.B. Maksić, M. Yáñez, J. Phys. Chem. A 109 (2005) 37.

[18] M. Rožman, D. Srzić, L. Klasinc, Int. J. Mass Spectrom. 253 (2006) 201.
[19] C. Bleiholder, S. Suhai, B. Paizs, J. Am. Soc. Mass Spectrom. 17 (2006) 1275.
[20] S. Gronert, D.C. Simpson, K.M. Conner, J. Am. Soc. Mass Spectrom. 20 (2009) 2116.
[21] G. Bouchoux, R.X. Chia, Croatica Chem. Acta 82 (2009) 47.
[22] G. Bouchoux, Mass Spectrom. Rev. 31 (2012) 391.
[23] A.F. Kuntz, A.W. Boynton, G.A. David, K.E. Colyer, J.C. Poutsma, J. Am. Soc. Mass Spectrom. 13 (2002) 72.
[24] O.E. Schroeder, E.J. Andriole, K.L. Carver, J.C. Poutsma, J. Phys. Chem. A 108 (2004) 326.
[25] E.J. Andriole, K.E. Colyer, E. Cornell, J.C. Poutsma, J. Phys. Chem. A 110 (2006) 11501.
[26] I.K. Webb, C.E. Muetterties, C.B. Platner, J.C. Poutsma, Int. J. Mass Spectrom. 316–318 (2012) 126.
[27] G. Chang, W.C. Guida, W. Clark Still, J. Am. Chem. Soc. 111 (1989) 4379.
[28] A.G. Baboul, L.A. Curtiss, P.C. Redfern, K. Raghavachari, J. Chem. Phys. 110 (1999) 7650.
[29] M.J. Frisch, Gaussian 03 & 09, Gaussian, Inc., Pittsburgh, PA, 2003 and 2009.
[30] E. Tarcsa, L.N. Marekov, G. Mei, G. Melino, S.C. Lee, P.M. Steinert, J. Biol. Chem. 271 (1996) 30709.
[31] A.K. Bence, P.A. Crooks, J. Enzyme Inhib, Med. Chem. 18 (2003) 383.
[32] H.M. Verheij, J.J. Volwerk, E.H.J.M. Jamen, W.C. Puyk, B.W. Dijkstra, J. Drenth, G.H. de Haas, Biochemistry 19 (1980) 743.
[33] D.R. Eyre, M.J. Glimcher, Proc. Nat. Acad. Sci. USA 69 (1972) 2594.