

REFERENCES

- May M, Abrams JA. Emerging insights into the Esophageal Microbiome. *Curr Treat Option Gastroenterol*. 2018;16:72-85.
- Muir AB, Benitez AJ, Dods K, Spergel JM, Fillon SA. Microbiome and its impact on gastrointestinal atopy. *Allergy*. 2016;71:1256-1263.
- David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559-563.
- Dawson DR 3rd, Branch-Mays G, Gonzalez OA, Ebersole JL. Dietary modulation of the inflammatory cascade. *Periodontol* 2000. 2014;64:161-197.
- EFSA NDA Panel (EFSA Panel on Dietetic Products Nutrition and Allergies). Guidance on the scientific requirements for health claims related to the immune system, the gastrointestinal tract and defence against pathogenic microorganisms. *EFSA J*. 2016;14:4369.
- Warners MJ, Vlieg-Boerstra BJ, Verheij J, et al. Elemental diet decreases inflammation and improves symptoms in adult eosinophilic oesophagitis patients. *Aliment Pharmacol Ther*. 2017;45:777-787.
- Van Rhijn BD, Vlieg-Boerstra BJ, Versteeg SA, et al. Evaluation of allergen-microarray-guided dietary intervention as treatment of eosinophilic esophagitis. *J Allergy Clin Immunol Pract*. 2015;136:1095-1097.
- De Vos WM, Castenmiller JJ, Hamer RJ, Brummer RJ. Nutridynamics—studying the dynamics of food components in products and in the consumer. *Curr Opin Biotechnol*. 2006;17:217-225.
- Wood LG, Shivappa N, Berthon BS, Gibson PG, Hebert JR. Dietary inflammatory index is related to asthma risk, lung function and systemic inflammation in asthma. *Clin Exp Allergy*. 2015;45:177-183.
- Van den Elsen L, Garssen J, Willemsen L. Long chain N-3 polyunsaturated fatty acids in the prevention of allergic and cardiovascular disease. *Curr Pharm Des*. 2012;18:2375-2392.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

DOI: 10.1111/all.13696

Three-dimensional structure of the wheat β -amylase Tri a 17, a clinically relevant food allergen

To the Editor,

Wheat is one of the most important staple foods worldwide and has been recognized as a potent food allergen source. Allergies are on the rise in western countries and the prevalence of food allergy has reached approximately 7% in children in the United States,¹ making improved diagnostics an important goal. Not all known potential allergens of wheat have been characterized so far and the prevalence of true wheat allergy is difficult to determine due to IgE cross-reactivity with grass pollen allergens.²

The wheat β -amylase (Tri a 17) has been found to bind IgE of wheat allergic patients,³ but its structure and allergenic activity have not been studied. To evaluate its possible clinical relevance and to shed some light on its biochemical properties, we elucidated the three-dimensional structure, measured the enzymatic activity, IgE-binding capacity, and allergenic activity of the recombinant enzyme.

The β -amylases of crop plants like wheat, soy, or barley have been under investigation for decades and the presence of two general forms has been described. One protein isoform found in all parts of the plant, termed ubiquitous β -amylase, and one variant specific to the endosperm, featuring an elongated C-terminal tail region are known. To date, only the ubiquitous form of wheat β -amylase has been sequenced.⁴

We screened a wheat seed cDNA library with serum IgE antibodies from patients suffering from wheat-induced food allergy and identified an IgE-reactive cDNA clone which was homologous to the

barley endosperm β -amylase at its C-terminus and to the sequence of wheat ubiquitous β -amylase at its N-terminus.

We then expressed a recombinant protein consisting of the ubiquitous wheat β -amylase sequence fused with the recovered IgE-reactive C-terminus (named Tri a 17_clone) and the unaltered ubiquitous β -amylase (termed Tri a 17_inactive) both as inclusion bodies in *Escherichia coli*. A refolding attempt yielded soluble, yet misfolded and aggregated protein. However, we were able to optimize the expression conditions to obtain natively folded amylase (named Tri a 17_active) (Figure S1 and S2 and the Methods section in this article's Online Repository).

The far-UV CD spectra of the inactive proteins indicate the presence of β -sheets and random coil signal, while Tri a 17_active exhibits the high α -helical content of a TIM-barrel (Figure 1B).

Purified Tri a 17_active crystallized readily and the structure was solved by X-ray crystallography to a resolution of 2 Å (PDB: 6GER). The structure superimposes well with the previously published structure of barley β -amylase (rmsd = 0.62 Å when 442 of 486 C α atoms are aligned with the PDB structure 2XFF⁵), which was used as the template for molecular replacement.

The crystal structure of Tri a 17_active shows the same (β/α)₈-barrel architecture found for other plant and bacterial β -amylases (Figure 1A).

Figure 1C shows the variation in the relative enzymatic activity of the β -amylase with respect to pH. The enzyme shows a tolerance

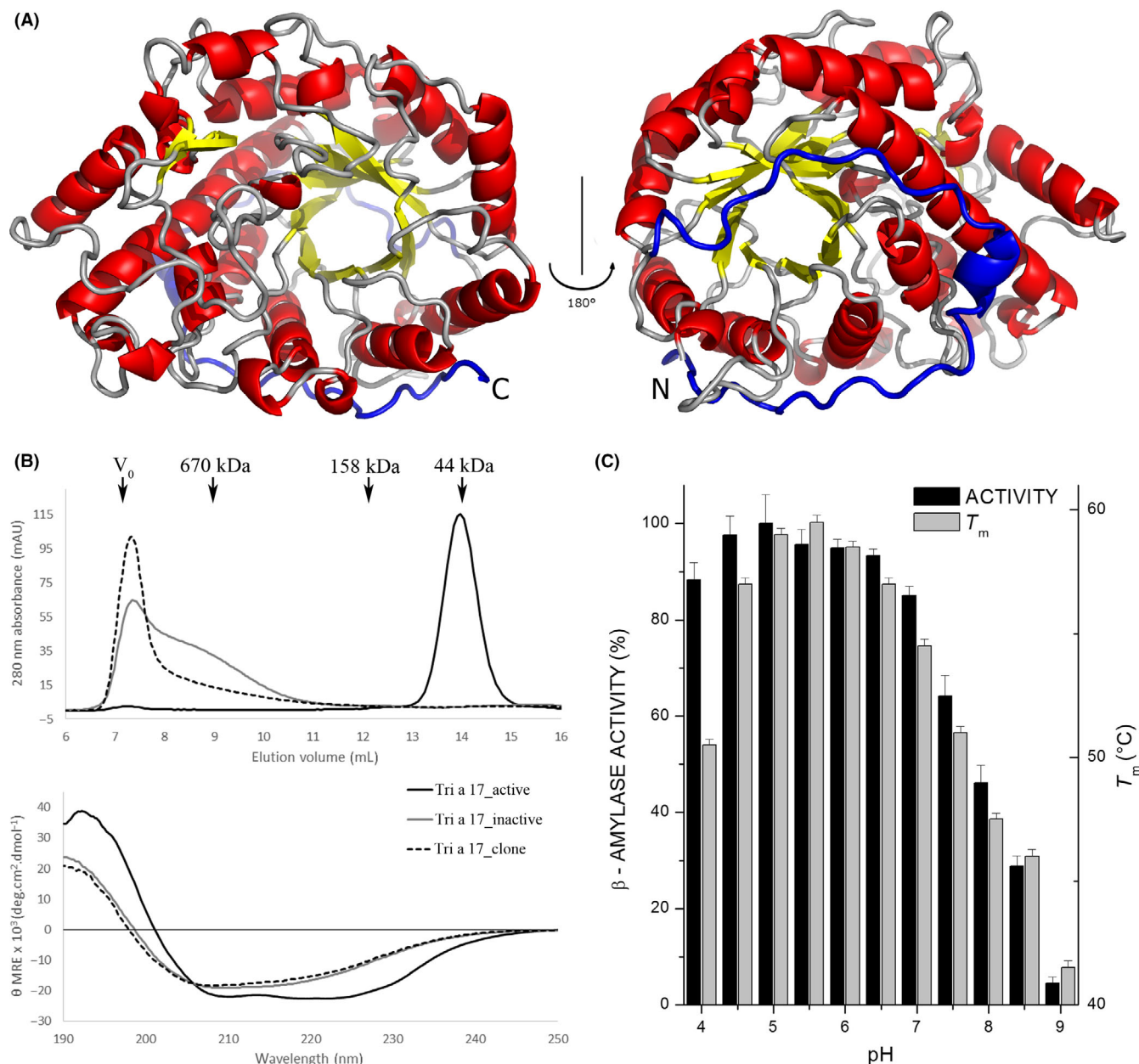


FIGURE 1 A, Cartoon model of Tri a 17 showing the $(\beta/\alpha)_8$ -barrel architecture with its C-terminal portion, which corresponds to the IgE-reactive peptide, in blue. B, Size exclusion chromatography traces of the recombinant amylases, marker protein elution volumes, and void volume are shown with arrows (top). Circular dichroism analysis of recombinant β -amylases. The spectra are expressed as mean residue ellipticities (θ -MRE) (y-axis) at given wavelengths (x-axis) (bottom). C, Percent maximum activity of Tri a 17_active as a function of pH (black). Tri a 17_active melting temperatures (T_m) against pH as measured by DSF (gray). Data are averages of three independent measurements with error bars showing the standard deviation

for low pH, showing maximum activity at pH 5 and retaining over 80% of its maximum activity even at pH 4. However, at higher pH values, a marked decrease in activity is seen, with more than 50% activity lost at pH 8. In general, Tri a 17 is highly active in a wide range of acidic pH conditions (4.0–7.0), similar to other β -amylases, such as the major β -amylases of barley.⁶

Figure 1C also shows the melting temperature of Tri a 17_active at different pH values, as measured by differential scanning fluorimetry. The protein is most stable in slightly acidic conditions. A maximum in melting temperature is seen at pH 5.5 with 59°C, with a tail

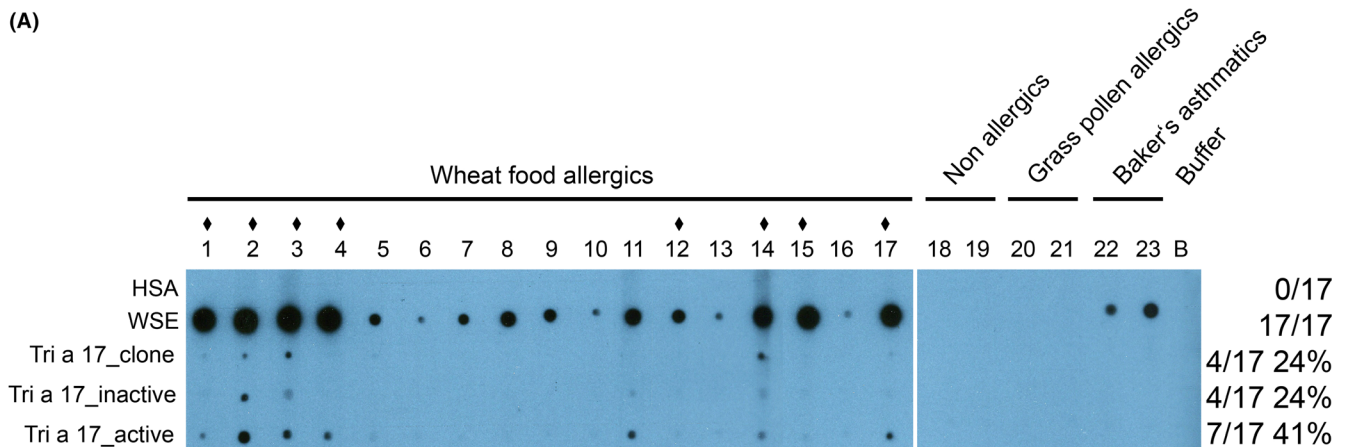
off to lower values as pH increases and a sharp drop in stability at pH 4.

The IgE reactivity of the three forms of Tri a 17 was assessed in non-denaturing RAST-based IgE dot blot experiments with sera from 17 wheat food allergic patients. Tri a 17_clone and Tri a 17_inactive were recognized by 24% (4 of 17) of wheat food allergic patients with varying intensities, whereas the folded enzyme (Tri a 17_active) was recognized by 41% (7 of 17) of wheat food allergic patients, indicating the presence of conformational IgE epitopes as well as linear epitopes. Since the enzymatic activity depends on the correct

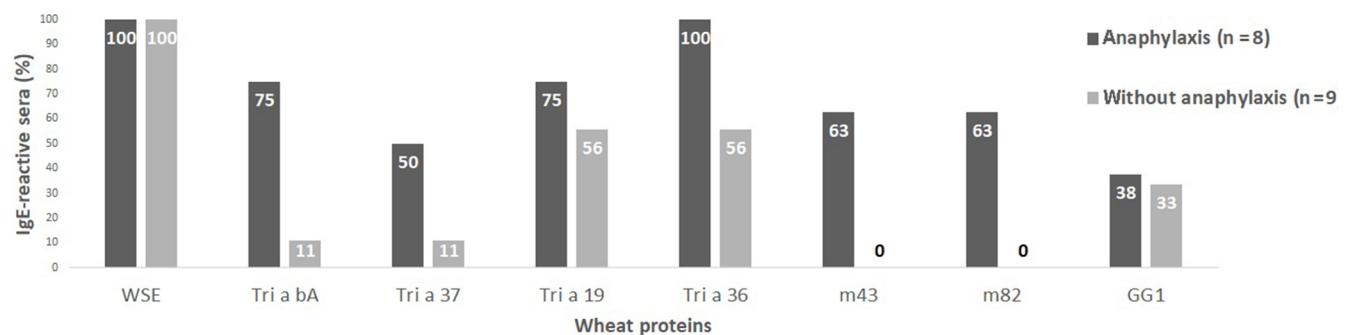
fold, there is an apparent correlation between enzymatic and allergenic activity of the β -amylase. IgE reactivity was specific to patients suffering from wheat food allergy. Non-allergic individuals (patients 18 and 19), grass pollen allergic patients (patients 20 and 21), and baker's asthma patients (patients 22 and 23) did not exhibit IgE reactivity to either form of beta amylase (Figure 2A).

Interestingly, 8 out of 17 wheat food allergic patients (47%) had a history of wheat-induced anaphylaxis. Of those, six were among the seven β -amylase positive patients. This correlation between anaphylaxis and Tri a 17 IgE recognition is statistically significant ($P = 0.015$). The relative risk of developing wheat-induced anaphylaxis, estimated by logistic regression, indicates a 24-fold higher

(A)



(B)



(C)

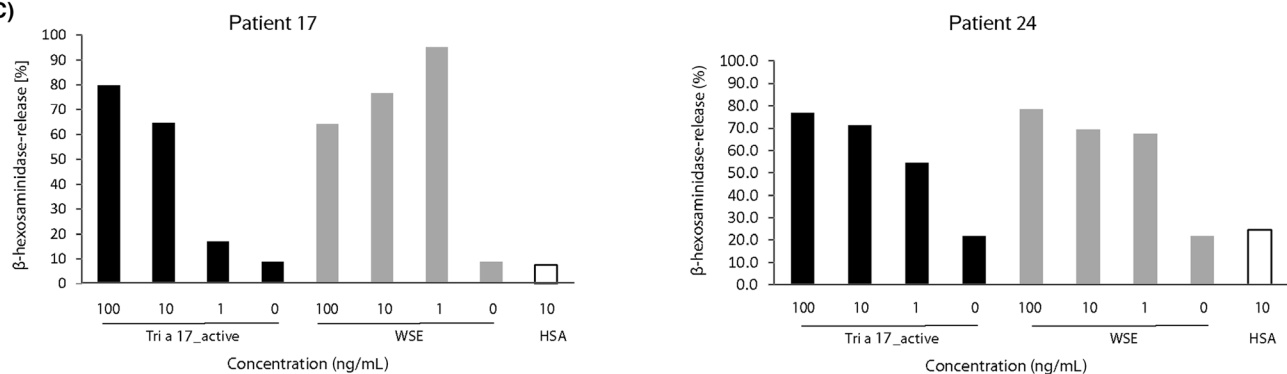


FIGURE 2 A, IgE reactivity of wheat β -amylases (Tri a 17_clone, Tri a 17_inactive and Tri a 17_active). Nitrocellulose-dotted human serum albumin (HSA), wheat seed extract (WSE), Tri a 17_clone and Tri a 17, in its inactive and active form, were tested with sera from wheat food allergic patients (1-17). For control purposes, sera from non-allergic individuals (18 and 19), grass pollen allergic patients (20 and 21) baker's asthma patients (22 and 23), or buffer alone (B) were included. Patients with a history of wheat-induced anaphylaxis are indicated by $\blacklozenge$ above their number. B, IgE reactivity of wheat food allergic patients with and without anaphylaxis. IgE-binding prevalences (y-axis: percentage of IgE-reactive sera) to wheat seed extract and wheat proteins (Tri a 17, Tri a 37, Tri a 19, Tri a 36, m43, m82, GG1) for patients with (black bars) and without anaphylaxis (gray bars). C, Allergenic activity of recombinant β -amylase Tri a 17_active and wheat seed extract. RBL cells were loaded with serum IgE from Tri a 17_active-reactive patients and incubated with increasing concentrations of Tri a 17_active (black) or WSE (gray) (100, 10, 1 ng/mL), buffer alone (0 ng/mL) or HSA (white) (10 ng/mL). β -Hexosaminidase releases are displayed as percentages of total β -hexosaminidase release on the y-axes

probability for β -amylase-reactive patients (Methods, Online Repository). Figure 2B shows the IgE recognition frequencies for the patients with and without anaphylaxis for β -amylase (Tri a 17) and other wheat allergens. High molecular weight glutenins (m43, m82,⁷ $P = 0.009$), low molecular weight glutenin (Tri a 36,⁸ $P = 0.08$), and alpha purothionin (Tri a 37,⁹ $P = 0.13$) also discriminate between patients with and without anaphylaxis, to varying degrees. However, omega 5 gliadin (Tri a 19, $P = 0.62$), gamma gliadin (GG1¹⁰), and wheat seed extract do not discriminate between patients with and without anaphylaxis.

Tri a 17_{active} was able to induce effector cell degranulation in basophil degranulation assays. Human Fc ϵ RI-expressing RBL cells were passively sensitized with sera from wheat food allergic patients which showed IgE reactivity to Tri a 17 (Figure 2A, Figure S3). Subsequent incubation with increasing concentrations of Tri a 17_{active} or wheat seed extract (WSE) showed release for both wheat food allergic patients tested (Figure 2B).

Within a barley kernel, β -amylases can reach 1%-2% of the total protein in the starchy endosperm.⁶ Assuming similarly high expression in wheat, amylase concentrations that showed effector cell release (1 μ g/L) are likely to be found in all flour-containing foodstuffs.

The multiple sequence alignment of β -amylase showed that highly homologous proteins occur in various plant species (Figure S2 in this article's Online Repository). Cross-reactive antigens were detected in rye, maize, oat, spelt, barley, soy, sunflower, and rice using rabbit sera.

The protein was detected in all processed cereal products (ie, brown bread, rye bread, and rolls) including gluten-free bread and even after baking, showing a remarkable persistence which may explain why it can induce severe reactions (Figure S4 in this article's Online Repository).

Based on our findings, wheat beta amylase has been given the official name "Tri a 17" by the WHO/IUIS allergen nomenclature subcommittee.

In summary, wheat β -amylase can be classified as a class I food allergen that sensitizes via the gastrointestinal tract.

Wheat β -amylase seems to be associated with severe allergic reactions upon wheat ingestion in sensitized patients and should be included in the panel of potential diagnostic marker molecules for severe wheat food allergy.

ACKNOWLEDGMENTS

Financial support by projects F4605 and F4604 of the Austrian Science Fund (FWF) is gratefully acknowledged. The authors gratefully acknowledge the support for installing the CD-core facility from NAWI Graz and the support of the beamline staff at the Elettra synchrotron source (Trieste, Italy) during data collection. Gerhard Hofer would like to acknowledge the financial support by the "Förderung Wissenschaftlicher Nachwuchs" program of the University of Graz. Sandra Wieser is a recipient of the EFIS-Acteria Doctoral Prize in Allergy sponsored by the Acteria foundation (Zurich,

Switzerland). Rudolf Valenta is recipient of a Megagrant of the Government of the Russian Federation, grant number 14.W03.31.0024.

CONFLICT OF INTEREST

Rudolf Valenta has received research grants from Biomay, Vienna, Austria and Viravaxx, Vienna, Austria and serves as a consultant for both companies. The other authors declare no conflicts of interest.

ORCID

Gerhard Hofer  <https://orcid.org/0000-0002-9248-6989>

Michael K. Bogdos  <https://orcid.org/0000-0003-1826-2706>

Ryosuke Nakamura  <https://orcid.org/0000-0002-7756-3798>

Nikolaos Papadopoulos  <https://orcid.org/0000-0002-4448-3468>

Rudolf Valenta  <https://orcid.org/0000-0001-5944-3365>

Walter Keller  <https://orcid.org/0000-0002-2261-958X>

Gerhard Hofer¹ 

Sandra Wieser²

Michael K. Bogdos¹ 

Pia Gattinger²

Ryosuke Nakamura³ 

Motohiro Ebisawa⁴

Mika Mäkelä⁵

Nikolaos Papadopoulos⁶ 

Rudolf Valenta^{2,7,8} 

Walter Keller¹ 

¹Institute of Molecular Biosciences, BioTechMed Graz, University of Graz, Graz, Austria

²Division of Immunopathology, Department of Pathophysiology and Allergy Research, Center for Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria

³Division of Medicinal Safety Science, National Institute of Health Sciences, Kanagawa, Japan

⁴Clinical Research Center for Allergology and Rheumatology, National Hospital Organization, Sagami National Hospital, Kanagawa, Japan

⁵Department of Allergy, Skin and Allergy Hospital, Helsinki University Central Hospital, Helsinki, Finland

⁶Allergy Department, 2nd Pediatric Clinic, University of Athens, Athens, Greece

⁷NRC Institute of Immunology FMBA of Russia, Moscow, Russia

⁸Laboratory for Immunopathology, Department of Clinical Immunology and Allergy, Sechenov First Moscow State Medical University, Moscow, Russia

Correspondence

Walter Keller, Institute of Molecular Biosciences, BioTechMed Graz, University of Graz, Graz, Austria.

Email: walter.keller@uni-graz.at

Wieser and Hofer contributed equally to the work.

REFERENCES

1. Yu W, Freeland DMH, Nadeau KC. Food allergy: immune mechanisms, diagnosis and immunotherapy. *Nat Rev Immunol*. 2016;16:751-765.
2. Constantin C, Quirce S, Poorafshar M, et al. Micro-arrayed wheat seed and grass pollen allergens for component-resolved diagnosis. *Allergy*. 2009;64:1030-1037.
3. Šotkovský P, Sklenář J, Halada P, et al. A new approach to the isolation and characterization of wheat flour allergens. *Clin Exp Allergy*. 2011;41:1031-1043.
4. Wagner G, KP Haeger PZ. The electronic plant gene register. *Plant Physiol*. 1996;112:1735-1736.
5. Rejzek M, Stevenson CE, Southard AM, et al. Chemical genetics and cereal starch metabolism: structural basis of the non-covalent and covalent inhibition of barley β -amylase. *Mol Biosyst*. 2011;7:718-730.
6. Lundgard R, Svensson B. The four major forms of barley β -amylase. Purification, characterization and structural relationship. *Carlsberg Res Commun*. 1987;52:313-326.
7. Baar A, Pahr S, Constantin C, et al. The high molecular weight glutenin subunit Bx7 allergen from wheat contains repetitive IgE epitopes. *Allergy*. 2014;69:1316-23.
8. Baar A, Pahr S, Constantin C, et al. Molecular and immunological characterization of Tri a 36, a low molecular weight glutenin, as a novel major wheat food allergen. *J Immunol*. 2012;189:3018-25.
9. Pahr S, Constantin C, Papadopoulos NG, et al. α -Purothionin, a new wheat allergen associated with severe allergy. *J Allergy Clin Immunol*. 2013;132:1000-3.
10. Srinivasan B, Focke-Tejkl M, Weber M, et al. Usefulness of recombinant γ -gliadin 1 for identifying patients with celiac disease and monitoring adherence to a gluten-free diet. *J Allergy Clin Immunol*. 2015;136:1607-1618.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

DOI: 10.1111/all.13699

Genotype-phenotype correlations in Brazilian patients with hereditary angioedema due to C1 inhibitor deficiency

To the Editor,

Hereditary angioedema (HAE) is a rare disease with autosomal dominant inheritance that affects 1 in every 50 000 individuals.¹ Patients with HAE present recurrent episodes of edema of subcutaneous tissue and submucosa that mainly affects the skin, gastrointestinal tract, and upper airways. In most cases, the disease results from the deficiency of C1 inhibitor (C1-INH) owing to mutations in *SERPING1*, the gene encoding C1-INH protein. The decrease in C1-INH activity leads to uncontrolled activation of the kallikrein-kinin system and increased formation of bradykinin, resulting in angioedema. HAE with normal C1-INH has also been characterized.² Diagnosis of C1-INH-HAE is based on clinical symptoms, positive family history, low levels and/or functional activity of C1-INH, and decreased C4.¹ Whether genetic analysis should be performed in routine clinical practice is yet debated. Variability in clinical presentation of HAE has prompted researchers to look for novel biomarkers. Kaplan and Maas have recently discussed the role of potential biomarkers, including blood levels of bradykinin or the pentapeptide Arg-Pro-Pro-Gly-Phe derived from bradykinin degradation and cleaved high-molecular-weight kininogen, in assessing HAE severity and response to therapeutic agents.³ The type of mutation in *SERPING1* could account for clinical phenotypes. Patients with missense mutations have been shown to present symptoms at later ages, milder clinical course, and lesser need for prophylactic medications than those with mutations that cause more profound changes in the molecule.^{4,5} Grouping patients with missense mutations in *SERPING1* affecting Arg466 at the reactive center with those carrying mutations leading to more significant changes in C1-INH molecule revealed association with more severe disease.^{6,7} In the present study, we aimed to identify

and characterize mutations in *SERPING1* in Brazilian patients with C1-INH-HAE and investigate the impact of the type of mutation on clinical features of the disease.

Sixty patients with C1-INH-HAE from 16 distinct families were characterized based on mutations in *SERPING1*. Diagnosis was established according to consensus criteria (Data S1).¹ One of the families (Family 7) has been previously reported by our group; we provide follow-up information, including five new members with HAE-1.⁸ Patients were divided into the following groups: group 1 comprising patients with deletion, insertion, duplication, or nonsense mutation and mutations affecting Arg466 at the reactive center ($n = 48$); and group 2 comprising patients with missense mutations, with the exception of mutations at Arg466 ($n = 12$). The rationale for dividing patients into these groups was the fact that mutations causing more profound alterations in the structure of the protein could lead to more severe disease, as previously reported.^{6,7}

Genomic DNA was extracted from whole blood or oral mucosa material using the DNA Wizard Genomic DNA Purification Kit (Promega, Madison, WI). PCR was conducted using *SERPING1* primers (Data S2, Table S1), and DNA sequences were analyzed using the SeqMan softwareTM (Lasergene; DNA Star, Inc., Madison, WI). Sequence variations were identified by comparison with GenBank accession number NM_000062.2, X54486. Protein effect was defined using the mature protein sequence, including signal peptide. Criteria of the American College of Medical Genetics and Genomics (ACMG) were applied for classifying sequence variants. Predicted functional analysis of missense mutations was performed using the bioinformatics tools SIFT (http://sift.jcvi.org/www/SIFT_enst_submit.html), PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2/>), and MutationTaster,

FIGURES

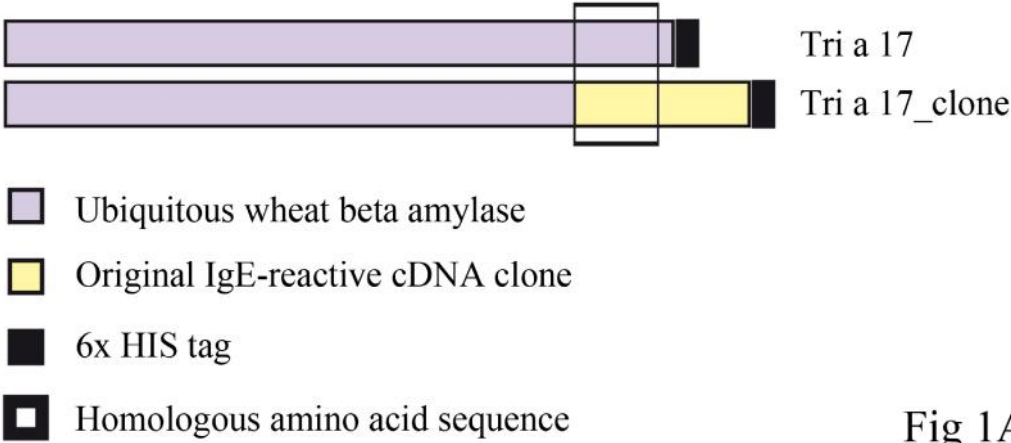


Fig 1A

Figure E1. Illustration of two constructs for gene synthesis of wheat beta amylase Tri a 17 and Tri a 17_clone. Construct “Tri a 17” codes for wheat beta amylase CAA67128.1 (violet). In the construct “Tri a 17_clone” the C-terminal part has been replaced by the beta amylase-derived original IgE-reactive cDNA clone (yellow). The boxed part highlights well aligning amino acid sequences. Both constructs contain a hexahistidine-tag, visualized in black.

Tri a 17 : -----* 20 * 40 * 60 * 80 * 100
Tri a 17 clone : -----MAGNMLANYVQVYVMLPLDVVSVDNKFEGKDEIRAQLKKLT-EAGVDGVMIDVWVGLVEGKGPAYDWSAYKQVFDLVHEAGLKLQAIMSP
Barley endo : -----A..N.R..L..R..V..V..L.E..QK..
Barley ubi : -----HV..YF..I..H.TP.NH..Y..D.R.Q..
Rye : -----
Rice : -----L..N..T..V..T..V..GS..E..L.R..Q..
Millet : -----Y..IIT..T..E..T..-A..E.GV..K..Q..
Maize : -----A..IT..T..E..T..-A..E.GV..R..K..Q..
Soy : -----MATSDS..L..P..G..N..V..DP.GLKE..LQ.R-A..V..II.L..Q..R..RSL.Q..Q.C..T..
Chestnut : -----MQASSLTINEK..P..G..TIN.VL.DKAG.EK..E.R-A..V..II.SQ..Q..RSL.QI.QDC..
Orange : -----MQASPAALTYDEK..P..PI..G..ITN..VL.DK.KLES..E.K-A..V..II.S..Q..RSL.E.IQQYE..
A. thaliana : -----MATNYNEKL..L..P..G..N..E..V..ADPETLET..R..KE..I..V..II.S..Q..T..TL.Q.IARL..I..
Radish : -----MAANYNEKL..L..P..G..D..E..V..VDPETLET..R..KE..I..V..II.S..Q..T..TL.Q.IARL..I..
Sw. potatoe : -----MAPIPGVMPIG..SL..G..NA..V..PDKEKVEDE..QVK-AG..C..V..II.A..Q..REL.Q..KKC..I..

Tri a 17 : -----* 120 * 140 * 160 * 180 * 200
Tri a 17 clone : -----HQCNGNVGDVVNIPIQVWRDVGATDPDIFYTNRGGRNIEYLTGLVDDQPLFHGRGTAVQMYADYMASFRENMKKFLDAGTIVDIEVGLGPAGEMRYPY
Barley endo : -----A..TR..DGH..N..S..T..D..V..
Barley ubi : -----LF..R..D..H..
Rye : -----G.TF..S..
Rice : -----I..I..SD..A..I..K..AE..T..V..
Millet : -----G..EDN..E..V..I..L..K..AD..V..
Maize : -----KSN..S..LT..I..L..K..AD..V..
Soy : -----I..L..I..ESNH..S..K..V..NE..I..IKI..S..K..SD..ES..L..I..
Chestnut : -----L..L..I..KS..V..LCN..K..S..NE..Y..EI..G..K..S..SD..FED..L..I..
Orange : -----T..K..LEI..E..N..S..N..K..I..HK..I..EI..S..R..SD..E..V..I..
A. thaliana : -----I..T..DN..Y..K..D..SI..NL..A..L..S..S..K..ADLIE..V..
Radish : -----T..K..E..ES..Y..R..D..S..NL..A..P..S..S..K..LEL..E..
Sw. potatoe : -----A..F..ILQI.DKN..A..N..Q..S..N..R..Q..LE..R..F..E..D..AD..K..D..C..A..L..

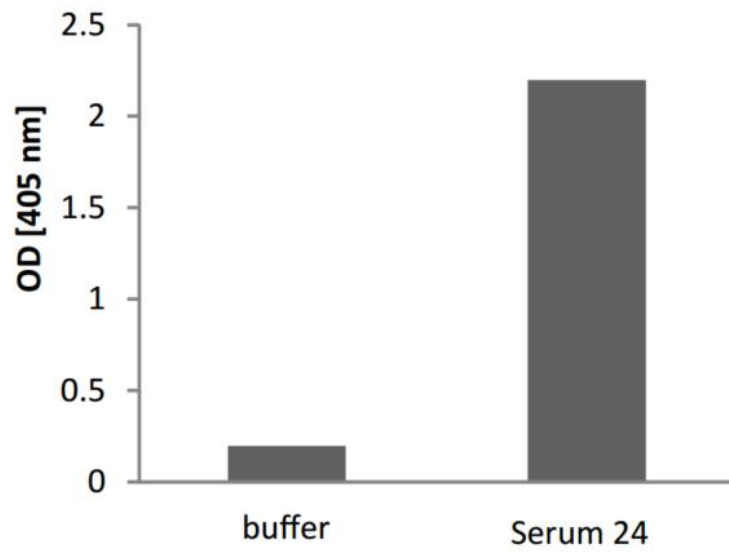
Tri a 17 : -----* 220 * 240 * 260 * 280 * 300
Tri a 17 clone : -----PQSQGWVFPFGIGEFICYDKYLEADFKAAAAKAGHPWELPDD-AGEYNDTPEKTQFFKNGTYLTKGKFFLSWYSNKLKHGDKILDEANKVFLGCRVQ
Barley endo : -----H..S..Q..AV..F..N..V..Q..R..R..S..R..A..N..R..YK..
Barley ubi : -----E..
Rye : -----G..E..
Rice : -----E..E..R..A..V..T..
Millet : -----Y..K..TA..D..A..Q..D..T..
Maize : -----V..Q..Q..EE..D..T..A..Q..D..T..
Soy : -----E..Q..K..V..R..K..V..S..G..S..V..T..LN..Q..A..K..K..
Chestnut : -----Q..K..E..E..TSV..N..T..TS..E..GQS..T..LS..Q..A..K..K..
Orange : -----E..Q..K..E..E..TAS..N..T..K..S..E..T..LF..E..N..A..K..K..
A. thaliana : -----Q..K..E..E..D..E..K..E..G..KD..VS..MT..F..Q..G..I..A..LK..N..
Radish : -----Q..K..E..E..D..E..K..E..G..RT..VS..E..T..T..F..Q..IG..I..A..LK..N..
Sw. potatoe : -----ET..Q..MV..W..E..VKQ..NAD..M..GKGT..T..D..E..RP..K..DM..T..I..QV..E..V..L..N..

Tri a 17 : -----* 320 * 340 * 360 * 380 * 400
Tri a 17 clone : -----LAIKISGIHWWYRVPNHAAELTAGYNNLDDRGYRTIARMLTRHSHASMNFTCAEMRDSQSEEAASAPEELVQQVLSAGVREGHLVACENALGRYDATAY
Barley endo : -----K..S..H..K..R..I..SQ..M..N..P..P..
Barley ubi : -----V..
Rye : -----N..R..CV..S..
Millet : -----V..T..H..P..N..S..
Maize : -----V..N..H..R..S..NL..N..
Soy : -----V..K..E..N..P..S..IL..L..PSD..G..Q..G..DIR..G..P..
Chestnut : -----A..V..K..AD..K..PV..I..S..Y..IL..L..SD..G..Q..G..N..E..G..P..RD..
Orange : -----A..V..LAD..N..S..I..S..YGIL..L..DAA..CG..Q..G..NIE..G..S..
A. thaliana : -----A..V..L..NHHS..FK..P..SK..YGIL..L..K..TDNTA..L..Q..E..KA..K..IE..G..ET..G..KG..
Radish : -----A..V..L..NHHS..YE..P..SK..YGIL..L..K..TDNTA..M..Q..M..KS..K..IE..G..ET..G..TKG..
Sw. potatoe : -----I..A..V..NHVS..F..VAG..P..A..TL..L..PA..Q..S..K..YID..G..P..
Tri a 17 : -----* 420 * 440 * 460 * 480 * 500
Tri a 17 clone : -----NTILRNARPGINKNGPPEHKLFGFTYLRSLNELLEGQNYATFQTVEKMHANLGHDPDSD---PVAPLERSKPEMPIEMILKAAQPKLEPFPPDKNTD
Barley endo : -----H..QS..Q..V..VN..K..DR..PR..Y..M..P..G..IS..Q..Q..QEH..
Barley ubi : -----E..Q..Y..Q..A..N..K..L..
Rye : -----S..H..D..S..K..KR..VSATN..LE..Q..M..GK..Q..H..A..E..
Millet : -----Q..R..A..Y..V..D..F..E..T..K..RR..DFN..K..
Maize : -----Q..H..V..D..FQE..T..K..RR..DYN..N..A..I..E..EV..D..
Soy : -----Q..IL..Q..V..N..KLSM..V..DD..QKS..FNI..KK..L..DQDYCANPQKYNHAI..KP..A..KI..VL..E..TK..TL..WLPE..
Chestnut : -----Q..L..N..V..E..KLRMY..V..DD..QEN..FNI..K..K..DQDYC..DPEKYSHHIG..K..YL..E..TE..M..W..E..
Orange : -----Q..L..N..V..E..KLRMY..V..DD..AEN..FKI..KI..K..DQDYC..DPOKYNHEID..LKF..SN..ELME..TKK..P..W..E..
A. thaliana : -----Q..L..N..V..P..K..KLRMY..DTVFQEN..FEL..KKL..R..DQDYCGDAKYGHEIV..KT..NSQLTL..D..AD..SGALKW..SE..
Radish : -----Q..L..N..V..HD..N..KLRMY..DTVFQEN..FQL..KK..R..DQD..CGDAEKYGHEIV..KTPNSHLT..D..AD..SGA..KW..SE..
Sw. potatoe : -----Q..L..V..N..V..L..KL..MS..L..DD..QTD..FEL..KK..K..D..DPS..NAI---SP..V..NSAIT..DELME..TKG..SR..S..WYDV..

Tri a 17 : -----* 520 * 540 * 560
Tri a 17 clone : -----LPVKDHTDVGDEVLPV----- :
Barley endo : -----QGLGGI..GGGEVED..TGGMGGEVQDDPTGGMGGEVQDDPTGGMGGEVEDPTGGMGGELEPPTV : 95%
Barley ubi : -----GPTGGM..GQ-AE-----GPTCGMGQV-KGPTGGMGQAEDPTSGMGG----- : 83%
Rye : -----N..V..A..L..AV----- : 94%
Rice : -----FL----- : 97%
Millet : -----M----- : 88%
Maize : ----- : 86%
Soy : -----MK..DG----- : 70%
Chestnut : -----MS..GGAL---SNLIDKIFSIFK----- : 67%
Orange : -----MN..GGTRGILAAALFGKIFSMPK----- : 65%
A. thaliana : -----K..DG----- : 64%
Radish : -----MK..DG----- : 63%
Sw. potatoe : -----M..DGSNPPD----- : 62%

Fig E2

Figure E2. Sequence comparison of wheat beta amylase (*Triticum aestivum*, accession number CAA67128.1) with homologous proteins from other plants: rye (*Secale cereale*, accession number CAA77817.1), barley endosperm-specific β -amylase (*Hordeum vulgare*, accession number AAO67356.1), barley ubiquitous β -amylase (*Hordeum vulgare*, accession number AAC64904.1), rice (*Oryza sativa*, accession number AAA33898.1), millet (*Sorghum bicolor*, accession number ADB81912.1), maize (*Zea mays*, accession number AAD15902.1), soy (*Glycine max*, accession number AAZ38831.1), chestnut (*Castanea crenata*, accession number AAK30294.1), orange (*Citrus trifoliata*, accession number AFQ33617.1), *Arabidopsis thaliana* (accession number BAA07842.1), radish (*Raphanus sativus*, accession number BAH20736.1), sweet potatoe (*Ipomoea batatas*, accession number BAA02286.1). The sequences of the recombinant proteins, Tri a 17 and Tri a 17_clone, are highlighted again in violet and yellow (original IgE-reactive cDNA clone). The percentages of sequence identities to wheat beta amylase are shown at the right margins of the sequences.



31
32

33 **Figure E3.** Specific IgE recognition of Tri a 17_active by patient #24. Serum from patient #24 and
34 buffer without addition of serum was tested for IgE reactivity to Tri a 17_active by ELISA. OD values
35 (y-axis) correspond to levels of bound IgE.

36

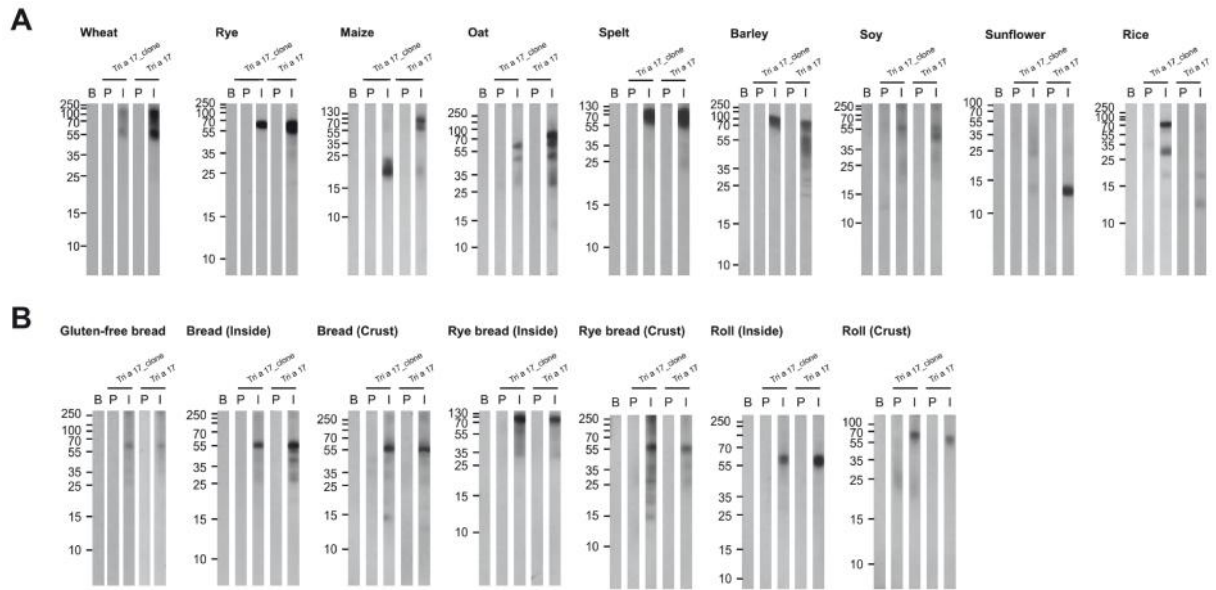


Figure E4. Detection of beta amylases (Tri a 17_clone, Tri a 17) in different cereals and types of bread. **(A)** Nitrocellulose-blotted extracts from cereals (wheat, rye, maize, oat, spelt, barley, soy, sunflower, rice) and **(B)** bread (gluten-free bread, inner part and crust of brown bread, rye bread and rolls) were incubated with Tri a 17_clone or Tri a 17_inactive specific rabbit antibodies (lanes: I), the corresponding pre-immune serum (lanes: P) or buffer as control. Molecular weights are indicated on the left side in kilo Dalton (kDa).

46 **Table S1**
47

Wavelength	0.9717
Resolution range	26.96 - 2.0 (2.072 - 2.0)
Space group	P 4 ₃
Unit cell	93.87 93.87 65.86 90 90 90
Total reflections	147815 (14537)
Unique reflections	38550 (3721)
Multiplicity	3.8 (3.8)
Completeness (%)	91.32 (97.05)
Mean I/sigma(I)	6.93 (3.76)
Wilson B-factor	15
R-merge	0.1203 (0.2861)
R-meas	0.1394 (0.3328)
R-pim	0.06928 (0.1676)
CC1/2	0.986 (0.906)
CC*	0.996 (0.975)
Reflections used in refinement	35919 (3723)
Reflections used for R-free	1253 (130)
R-work	0.1582 (0.1735)
R-free	0.1886 (0.1995)
CC(work)	0.964 (0.924)
CC(free)	0.958 (0.884)
Number of non-hydrogen atoms	4474
macromolecules	3950
solvent	524
Protein residues	488
RMS(bonds)	0.003
RMS(angles)	0.59
Ramachandran favored (%)	98.35
Ramachandran allowed (%)	1.44
Ramachandran outliers (%)	0.21
Rotamer outliers (%)	1.68
Clashscore	1.54
Average B-factor	20.96
macromolecules	19.82
solvent	29.5

Methods

Biological materials

Wheat seeds were obtained from AGES (Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH). Seeds from rye, maize, oat, spelt, barley, soy, sunflower and rice were bought at a local drugstore. Human serum albumin (HSA) was purchased from Behring (Marburg, Germany) and *E.coli* strain BL21 (DE3) from Stratagene (La Jolla, CA). Specific rabbit antibodies were raised by immunization of a rabbit each with purified recombinant beta amylase Tri a 17_inactive and Tri a 17_clone (200 µg per injection) using once Freund's complete adjuvant and twice Freund's incomplete adjuvant (Charles River, Kisslegg, Germany). Pre-immune sera were obtained from the rabbits before immunization.

Preparation of protein extracts and immunoblots

SDS-protein extracts for Western blotting (1) were prepared from equal amounts (weight) of seeds from wheat (cultivar Grandios), rye, maize, oat, spelt, barley, soy, sunflower and rice. Different sorts of bread (gluten-free bread, brown bread, rye bread, rolls) were bought at a local bakery and homogenized in a grinding machine (CTC, Clatronic International, Kempen, Germany). The powder obtained from 3 grams of tissue was extracted in 32 ml SDS-sample buffer by boiling for 10 minutes and centrifuged at 10 000 x g for 10 minutes at 4°C to obtain the cleared extracts. Aqueous wheat seed extract (WSE) was prepared as described.(2) The protein concentration was determined with the Micro BCA Protein Assay Kit (Pierce, Rockford, IL). Extracts were stored at -20°C until use.

Extracts were analyzed by SDS-PAGE with Coomassie-blue staining and blotted onto nitrocellulose. Nitrocellulose-blotted extracts were tested with beta amylase-specific rabbit antibodies, and for control purposes with the corresponding pre-immune serum (dilution 1:10.000) and buffer alone. Bound IgG antibodies were detected with ¹²⁵I-labeled goat anti-rabbit antibodies (Perkin Elmer, Boston USA) and visualized by autoradiography.

Patient's sera

Sera were obtained from 17 European wheat food allergic patients (Austria n=2, Finland n=7, Greece n=8). One patient (patient 24) from

Japan was additionally included for rat basophil leukemia assays. Patients were diagnosed on the basis of a case history demonstrating that allergic symptoms (systemic anaphylaxis, airway symptoms, gastrointestinal symptoms, itching of the mouth, skin symptoms) were unambiguously related to the ingestion of wheat or wheat products.

A grading of symptoms and definition of anaphylaxis was performed according to the international position paper and recommendations for the definition of anaphylaxis. (3)(4) In 8 of the patients, open food challenge was performed and positive challenge results were obtained in 7 patients. For patients from Finland (patients #5-9 and #11), the PRACTall consensus criteria were used according to Sampson *et al* to describe clinical symptoms (3). Patient #5 had an immediate reaction upon open, oral wheat food challenge, showing symptoms in the skin (urticaria) and in the respiratory tract such as cough and rhinitis. Patient #6 reacted immediately with skin manifestations, such as urticaria and erythema. Patient #7 had a delayed type systemic allergic reaction, displaying lip edema, rhinitis, sneezing and itching of the mouth. Patient #8 demonstrated immediate symptoms, such as globus sensation and itching of the mouth. Patient #9 reacted immediately with gastrointestinal symptoms, such as abdominal pain and nausea. Patient #11 had delayed symptoms affecting the skin, showing itching and exacerbation of atopic eczema. Challenge tests in Greek patients were considered positive if objective signs were present (3). Patient #17 had sneezing, rhinitis and urticaria. Skin prick tests with wheat seed extract were performed in 14 patients, and were positive in all of them. For each patient IgE-mediated sensitization to wheat was confirmed by measurements of allergen-specific IgE (CAP-FEIA, Thermofisher, Uppsala, Sweden). Studies were performed on anonymized sera with permission of the Ethics Committee of the Vienna General Hospital.

Expression and purification of recombinant wheat beta amylase in its active and inactive form

The cDNA coding for the mature wheat beta amylase (accession number CAA67128.1, designated Tri a 17) as well as the beta amylase whose C-terminal part has been replaced by the beta amylase-derived original IgE-reactive cDNA clone (Tri a 17_clone) with an additional 3' sequence coding for a hexahistidine tag were

produced as synthetic genes with codons optimized for expression in *E. coli* and subcloned into the pET17b expression vector (GenScript, NJ, USA). The beta amylases, expressed in *Escherichia coli* BL21 (DE3) were purified by nickel affinity chromatography from the soluble fraction (Quiagen, Hilden, Germany) as previously described.(5) Recombinant allergens were stored in 10 mM NaH₂PO₄ buffer pH 7.0 at -20°C. The protein concentrations were determined by BCA assay (Pierce, Rockford, IL) and purity was assessed by Coomassie Blue-stained 14% SDS-PAGE. Proteins were analysed under reducing and non-reducing conditions.(6) By the use of this method, we obtained enzymatically inactive recombinant beta amylases.

In parallel, we expressed an active beta amylase using another protocol in $\Delta trxB$, Δgor strain *E. coli* (SHuffle, New England Biolabs) starting with an overnight culture in 2% glucose LB medium at 30°C. Main cultures were inoculated by adding overnight culture to a final OD₆₀₀ of 0.1 into LB medium containing 0.5% glucose. Expression was induced by adding IPTG to a final concentration of 0.4 mM when the main cultures reached an OD₆₀₀ of 0.6-0.8. Cultures were then either left at 30°C for 4 hours or cooled to 16°C and shaken overnight.

The *E. coli* pellets were sonicated for lysis in a 100 mM MES pH 6.5, 300 mM NaCl and 50mM imidazole buffer. The soluble fraction after centrifugation at 15 000g for 30 min was applied to an Immobilized Metal Affinity Chromatography (IMAC) column (HisTrap 5ml FF). After washing with 10 column volumes of lysis/washing buffer, the protein was eluted with the same buffer with increased imidazole concentration (300 mM).

For crystallization, the protein was further purified by bringing the solution to 40 % (v/v) ethanol, removing the precipitated proteins by centrifugation and changing the buffer of the beta amylase-containing supernatant to 50 mM MES pH 6, 50 mM NaCl using a micro centrifugation filter and stored at 4°C. In this case, proteins were analysed using 12% or 18% SDS-PAGE stained with Coomassie Blue.

For the activity and stability assays, a size exclusion chromatographic purification in 100mM NaCl, 10mM MES pH 6.5 was performed. (Superdex 200 10/300).

Secondary structure determination

Circular dichroism analyses were performed with protein concentrations of 0.5 mg/ml Tri a

17 in 25mM MES pH 6 and 25mM NaCl. Samples were analyzed in 0.2mm cuvettes using a Jasco J-715 spectropolarimeter at room temperature. The spectra were measured from 260 to 190nm, 0.2 nm data pitch, 50 nm per minute and a response time of 2 sec. Secondary structure predictions were performed using Dichroweb (7) and the CDSSTR method.

Determination of enzymatic activity

To achieve a range of conditions with different pH values but the same buffer system, a multicomponent buffer system was used [1 : 2 : 2 molar ratio of L-malic acid, MES and Tris; 1M total concentration]. Mixing of two buffers of the same buffer system, but with different pH values, one with pH 4 and one with pH 9, gives a reasonably linear pH gradient.(8)

The enzymatic activity of Tri a 17_active was measured by a modified version of the 3,5-dinitrosalicylic acid assay described by Fischer and Stein.(9) The modified protocol is as follows:

Substrate solution (45 μ l, 1 % soluble starch [S9765, Sigma-Aldrich] in ddH₂O) mixed with multicomponent buffer (8) (10 μ l total) and protein solution (45 μ l, 2 mg/l in 1 mM MES, 100 mM NaCl, pH 6.5) were equilibrated to 25 °C. The two solutions were mixed and shaken at 450 rpm. After 7 minutes, the reaction was quenched by adding stopping reagent (100 μ l, 44 mM 3,5-dinitrosalicylic acid, 1.06 M sodium potassium tartrate, 0.4 M NaOH) and was subsequently placed at 100 °C for 5 minutes. Following, it was placed on ice for 3 minutes after which 1 ml of ddH₂O was added. Blank samples were created by adding the stopping reagent before the protein solution.

The absorbance of the solution at 540 nm was measured in a 1 cm cell, using a Hitachi U-2001 spectrophotometer (Hitachi High-Technologies Corporation, Tokyo, Japan). The blank corrected absorbance was used to calculate the number of moles of maltose released (reducing ends equivalents) by means of a maltose standard curve, created with D(+)-maltose (Carl Roth, Karlsruhe, Germany). The activity of the enzyme in units was then calculated.

One unit of activity was defined as the release of one micromole of maltose per minute at 25°C, under the conditions outlined above.

Thermostability

The effect of pH on the melting temperature of Tri a 17_active was investigated using differential scanning fluorimetry. Protein solution (10 μ l, 0.16 g/l in 10 mM MES, 100 mM NaCl), multicomponent buffer (8) (10 μ l total) and SYPRO orange (5 μ l, 1:200 prediluted in ddH₂O) were mixed in 96 well plates Using a C1000 Thermal Cycler (Bio Rad), the solution was heated 1°C per minute, from 25°C to 95°C. Fluorescence was measured every 0.5°C, using channel 2 of a CFX Real Time system (Bio Rad). The melting temperature T_m was calculated as the minimum of the first derivative of the fluorescence vs. the temperature.

Crystallization and structure refinement

For screening, the protein solution (in MES pH 6, 50 mM NaCl) was used in concentrations of 2-4 mg/ml. A mixture of 0.5 μ l protein with 0.5 μ l JCSG+ screen was performed in 96 well microbatch plates and covered with a combination of three parts paraffin to one part silicone oil.

The condition H3 (0.1M bis-Tris pH 5.5, 25% PEG 3350) was optimized and repeatable crystallization was achieved with 25% PEG 3350, 0.1M bis-Tris pH 5.0.

Molecular graphics and analyses were performed with PyMOL (The PyMOL Molecular Graphics System, Version 1.7.1 Schrödinger, LLC.).

The diffraction data were integrated using XDS (10), scaled and merged using SCALA of the CCP4 program suite (11). Molecular replacement was achieved using the PDB model 2XFR (12) of the barley amylase in the Phenix (Python-based Hierarchical ENvironment for Integrated Xtallography (13)) package, which was also used for refinement with alternating manual adaptation in Coot (14). The refined coordinates and structure factors have been deposited in the protein databank with the PDB-ID 6GER and experimental data are summarized in Table S1.

Dot-blotting

IgE-reactivity of wheat seed extract, beta amylases and additional wheat allergens was tested by non-denaturing, RAST-based IgE-dot blot experiments (5). Aliquots of each recombinant wheat allergens (0.5 μ g/dot) were dotted onto nitrocellulose strips (Schleicher & Schuell, Dassel, Germany). Aliquots of HSA (0.5 μ g/dot) and aqueous wheat seed extract (2 μ g/dot) were dotted for control purposes. The

strips were incubated with 1:10 diluted sera of 17 wheat food allergic patients, non-allergic individuals (n=2), grass pollen allergic patients (n=2), baker's asthma patients (n=2) and a buffer control. Bound IgE antibodies were detected with ¹²⁵I-labeled goat anti-human IgE antibodies (Demeditec Diagnostics, Kiel, Germany) and visualized by autoradiography (1). IgE-reactivity of omega-5-gliadin (f416: rTri a 19) was determined by ImmunoCAP measurements (ThermoScientific, Uppsala, Sweden) with a cut-off of 0.35 kUA/L.

BAT assay

Rat basophilic leukemia (RBL) cells expressing the $\alpha/\beta/\gamma$ subunits of the human Fc ϵ RI (15) were incubated with sera from patients with IgE reactivity to Tri a 17_active (patients 17 and 24 were chosen because they supplied sufficient serum for the experiment). Degranulation was induced by different concentrations of either Tri a 17_active, WSE or HSA as control. Released β -hexosaminidase was measured (16) and results are reported as percentage of total β -hexosaminidase release. All measurements were performed in triplicates.

ELISA

For certain patients, IgE reactivity to Tri a 17_active was tested by ELISA assay. Aliquots of 100 μ l Tri a 17_active (5 μ g/ml diluted in PBS) were coated over night at 4°C. After blocking for 3 hours with 3% BSA in PBST, serum (1:10 diluted) or buffer alone (negative control) was applied and incubated over night at 4°C. Bound IgE antibodies were detected as previously described.(17)

Association of wheat allergens with anaphylaxis

The odds ratio was calculated to determine the relative risk of a beta amylase positive patient to develop severe allergic symptoms. For this purpose, a cross tab was used to fulfill the current question. In this cross tab, eight patients had neither anaphylaxis nor IgE antibodies against beta amylase, whereas six patients have both. One patient had IgE but no anaphylaxis and two patients suffered from anaphylaxis but had no IgE antibodies. According to this result the relative risk was estimated by means of logistic regression as described (18).

The p-values were calculated by Fisher's exact test in R (19) using the data found in figure 2B.

Biopolymers 2008;**89**:392–400.

1. Valenta R, Duchene M, Ebner C, Valent P, Sillaber C, Deviller P et al. Profilins constitute a novel family of functional plant pan-allergens. *J Exp Med* 1992;**175**:377–385.
2. Baar A, Pahr S, Constantin C, Scheiblhofer S, Thalhamer J, Giavi S et al. Molecular and immunological characterization of Tri a 36, a low molecular weight glutenin, as a novel major wheat food allergen. *J Immunol* 2012;**189**:3018–3025.
3. Sampson HA, Gerth van Wijk R, Bindslev-Jensen C, Sicherer S, Teuber SS, Burks AW et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology–European Academy of Allergy and Clinical Immunology PRACTALL consensus report. *J Allergy Clin Immunol* 2012;**130**:1260–1274.
4. Sampson HA, Muñoz-Furlong A, Campbell RL, Adkinson NF, Allan Bock S, Branum A et al. Second symposium on the definition and management of anaphylaxis: Summary report - Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network Symposium. *Ann Emerg Med* 2006;**47**:373–380.
5. Pahr S, Constantin C, Mari A, Scheiblhofer S, Thalhamer J, Ebner C et al. Molecular characterization of wheat allergens specifically recognized by patients suffering from wheat-induced respiratory allergy. *Clin Exp Allergy* 2012;**42**:597–609.
6. Chen K-W, Blatt K, Thomas WR, Swoboda I, Valent P, Valenta R et al. Hypoallergenic Der p 1/Der p 2 combination vaccines for immunotherapy of house dust mite allergy. *J Allergy Clin Immunol* 2012;**130**:435–443.e4.
7. Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 2008;**89**:392–400.
8. Newman J. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr Sect D Biol Crystallogr* 2004;**60**:610–612.
9. Fischer EH. *Biochemical Preparations*. 1961
10. Kabsch W. Xds. *Acta Crystallogr Sect D Biol Crystallogr* 2010;**66**:125–132.
11. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 2011;**67**:235–242.
12. Rejzek M, Stevenson CE, Southard AM, Stanley D, Denyer K, Smith AM et al. Chemical genetics and cereal starch metabolism: structural basis of the non-covalent and covalent inhibition of barley β -amylase. *Mol Biosyst* 2011;**7**:718–730.
13. Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 2010;**66**:213–221.
14. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 2010;**66**:486–501.
15. Nakamura R, Uchida Y, Higuchi M, Nakamura R, Tsuge I, Urisu A et al. A convenient and sensitive allergy test: IgE crosslinking-induced luciferase expression in cultured mast cells. *Allergy* 2010;**65**:1266–1273.
16. Gieras A, Focke-Tejkl M, Ball T, Verdino P, Hartl A, Thalhamer J et al. Molecular determinants of allergen-induced effector cell degranulation. *J Allergy Clin Immunol* 2007;**119**:384–390.
17. Resch Y, Weghofer M, Seiberler S, Horak F, Scheiblhofer S, Linhart B et al. Molecular characterization of Der p 10: a diagnostic marker for broad sensitization in house dust mite allergy. *Clin Exp*

Allergy 2011;**41**:1468–1477.

18. Pahr S, Constantin C, Papadopoulos NG, Giavi S, Mäkelä M, Pelkonen A et al. α -Purothionin, a new wheat allergen associated with severe allergy. *J Allergy*

Clin Immunol 2013;**132**:1000–1003.e4.

19. R Development Core Team. R: A language and environment for statistical computing. 2015. doi:10.1007/978-3-540-74686-7



Preliminary Full wwPDB X-ray Structure Validation Report ⓘ

Apr 27, 2018 – 08:23 AM BST

Deposition ID : D_1200009740
PDB ID : *(not yet assigned)*

This is a Preliminary Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB Deposition System during initial deposition but before annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	rb-20031021
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031021

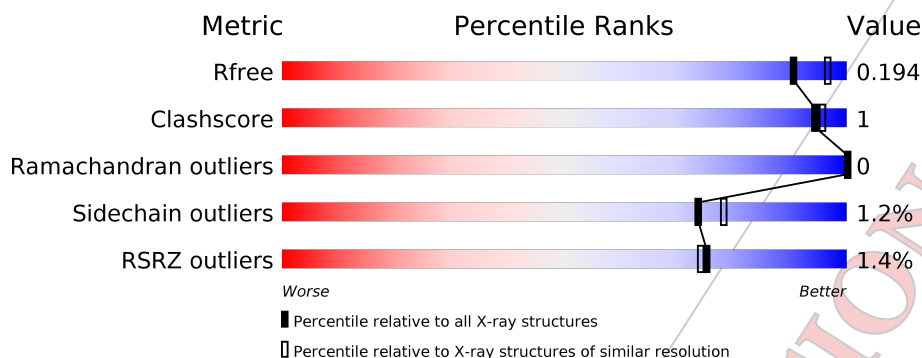
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	7193 (2.00-2.00)
Clashscore	122126	8267 (2.00-2.00)
Ramachandran outliers	120053	8166 (2.00-2.00)
Sidechain outliers	120020	8165 (2.00-2.00)
RSRZ outliers	108989	7011 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	509	92%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8214 atoms, of which 3749 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called beta-Amylase of wheat.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	488	7699	2523	3749	669	738	20	0	15	0

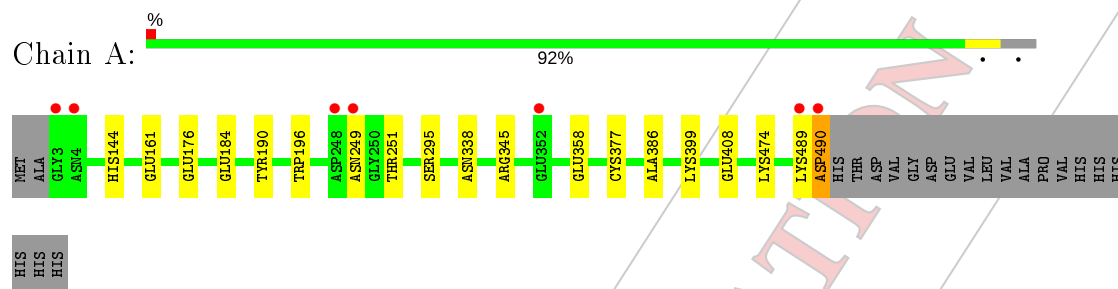
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	S	515	Total	O	0	0
			515	515		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: beta-Amylase of wheat



4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	93.87Å 93.87Å 65.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.06 – 2.00 27.06 – 1.70	Depositor EDS
% Data completeness (in resolution range)	91.5 (27.06-2.00) 79.0 (27.06-1.70)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.54 (at 1.70Å)	Xtriage
Refinement program	phenix.refine 1.11.1_2575, PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.157 , 0.193 0.158 , 0.194	Depositor DCC
R_{free} test set	1257 reflections (2.22%)	wwPDB-VP
Wilson B-factor (Å ²)	11.5	Xtriage
Anisotropy	0.589	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 43.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.045 for h,-k,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8214	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.25	0/4110	0.43	0/5570

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3950	3749	3803	11	0
2	S	515	0	0	5	0
All	All	4465	3749	3803	11	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (11) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:249:ASN:OD1	2:S:393:HOH:O	2.07	0.70
1:A:408:GLU:OE2	2:S:238:HOH:O	2.14	0.65
1:A:144:HIS:O	2:S:78:HOH:O	2.15	0.63
1:A:184:GLU:OE1	1:A:295[B]:SER:OG	2.18	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:176:GLU:OE2	1:A:338[B]:ASN:ND2	2.34	0.54
1:A:345[B]:ARG:NH1	2:S:305:HOH:O	2.46	0.48
1:A:474:LYS:NZ	2:S:270:HOH:O	2.49	0.45
1:A:345[B]:ARG:HD2	1:A:358:GLU:OE2	2.17	0.44
1:A:386:ALA:HB2	1:A:490:ASP:HA	1.98	0.44
1:A:190:TYR:HB2	1:A:196:TRP:CD2	2.55	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	500/509 (98%)	493 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	421/425 (99%)	415 (99%)	6 (1%)	69	74

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	161[A]	GLU
1	A	161[B]	GLU
1	A	377	CYS
1	A	399	LYS
1	A	489	LYS
1	A	490	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	194[B]	GLN
1	A	445	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	488/509 (95%)	-0.35	7 (1%) 75 74	6, 17, 38, 82	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	GLY	4.7
1	A	490	ASP	4.0
1	A	489	LYS	3.5
1	A	248	ASP	3.2
1	A	4	ASN	2.9
1	A	352	GLU	2.3
1	A	249	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.