The three-dimensional structure of human transaldolase

Stina Thorell[a], Peter Gergely Jr.[b], Katalin Banki[b], Andras Perl[b], Gunter Schneider[a,*]

[a] *Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden*
[b] *Department of Medicine and Microbiology and Immunology, College of Medicine, State University of New York, Syracuse, NY 13210, USA*

Received 22 March 2000; received in revised form 22 May 2000

Edited by Matti Saraste

Abstract The crystal structure of human transaldolase has been determined to 2.45 Å resolution. The enzyme folds into an α/β barrel structure and is thus similar in structure to other class I aldolases. Structure-based sequence alignment of available sequences of the transaldolase subfamily reveals that eight active site residues are invariant in the whole subfamily. Other invariant residues are mainly involved in the formation of the hydrophobic core of the enzyme. Noteworthy is a hydrophobic cluster consisting of five invariant residues. Human transaldolase has been implicated as an autoantigen in multiple sclerosis and four immunodominant peptide segments are located at the surface of the enzyme, accessible to autoantibodies.

Key words: Transaldolase; Class I aldolase; Protein crystallography; Multiple sclerosis

1. Introduction

Transaldolase (EC 2.2.1.2), one of the enzymes in the non-oxidative branch of the pentose phosphate pathway, catalyzes the reversible transfer of a dihydroxyacetone moiety from fructose 6-phosphate to erythrose 4-phosphate, giving sedoheptulose 7-phosphate and glyceraldehyde 3-phosphate. The reaction mechanism includes formation of a Schiff base intermediate by an active site lysine [1,2], a characteristic of class I aldolases [3]. Amino acid sequence comparisons suggest that transaldolases can be divided into two subclasses: the transaldolase and the MipB/TalC subfamilies. Most enzymes of the transaldolase subfamily consist of around 320–350 amino acids per polypeptide chain. The crystal structure analysis of *Escherichia coli* transaldolase, a representative of this subfamily, revealed that these enzymes contain an eight-stranded α/β barrel fold and that they are related to other class I aldolases by a circular permutation [4,5]. This permutation of the transaldolase gene shifts the catalytic lysine residue from the classical position in β-strand 6 in class I aldolases to β-strand 4 of the α/β barrel.

Multiple sclerosis (MS) is considered an autoimmune disease of the central nervous system resulting in chronic inflammation. Human transaldolase has been implicated as an autoantigen in this process based on detection of T-cell autoreactivity and autoantibodies in the serum and cerebrospinal fluid of MS patients [6–8]. These findings prompted us to initiate genetic, biochemical [9,10] and structural studies with the aim of characterizing the human enzyme in more detail. Here, we report the three-dimensional structure of human transaldolase at 2.45 Å resolution.

2. Materials and methods

2.1. Protein purification and crystallization

Human transaldolase was produced in *E. coli* and purified as described previously [10]. Crystals were obtained by the hanging drop method using a mother liquid consisting of 27% (w/v) PEG4000, 0.08% NaN_3, 0.3 M NH_4Ac, pH 4.4. For crystallization, 3 μl of a protein solution (6 mg/ml in 50 mM Tris, pH 7.0) was mixed with an equal amount of mother liquid and the droplets were then left to equilibrate with 1 ml of the mother solution at 20°C. Crystals of about 0.1 mm in length grew within 24 h.

2.2. Data collection and structure solution

Diffraction data (Table 1) were collected at 100 K to 2.45 Å resolution at beamline 711, MAX Laboratory, University of Lund, using a flash-frozen crystal soaked in 20% ethylene glycol as cryoprotectant.

Table 1
Data collection and refinement statistics

Data collection:	
Space group	$P2_1$
Cell dimensions (Å)	$a = 45.6$, $b = 113.6$, $c = 69.2$, $\beta = 101.4°$
Resolution[a] (Å)	23.96–2.45 (2.54–2.45)
Number of reflections	
Total	60 447
Unique	22 520
Completeness[a] (%)	88.2 (79.7)
I/σ[a]	7.4 (3.3)
R_{sym}[a]	0.080 (0.206)
Refinement:	
Resolution[a] (Å)	20–2.45 (2.60–2.45)
Number of reflections	
Total	22 371
working set	21 230
test set	1 141
Number of residues	322 × 2
Number of water molecules	92
R values[a] (%)	
Refinement	22.5 (24.7)
Free	25.8 (28.4)
Deviations from ideal (rmsd)	
bond distances (Å)	0.006
bond angles (°)	1.2
dihedrals (°)	20.6
impropers (°)	0.80
B values (Å^2)	
protein (all atoms)	21.4
solvent	16.2
Ramachandran plot	
% non-glycine or non-proline residues	
in most favored regions	92.8
additional allowed regions	6.8
generously allowed regions	0.3
disallowed regions	0.0

[a] Values in parentheses are for the highest resolution shell.

*Corresponding author. Fax: (46)-8-327626.
E-mail: gunter@alfa.mbb.ki.se

PII: S0014-5793(00)01658-6

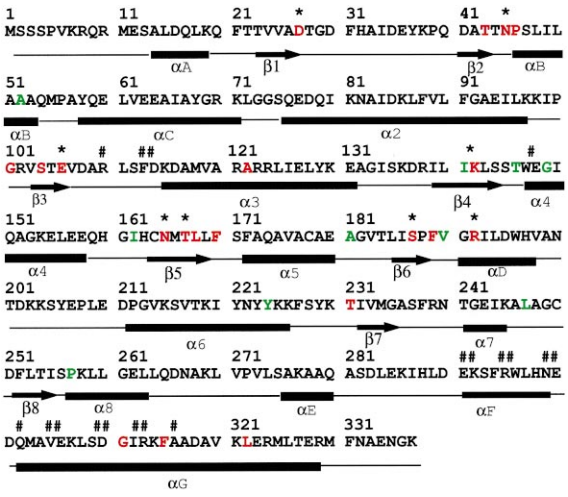

Fig. 1. Conservation of amino acid sequence in the transaldolase subfamily. The sequence of human transaldolase is shown together with the secondary structural elements. Invariant residues are shown in red, while residues in green are more than 90% conserved. Active site residues are marked by * and # denotes residues participating in the dimer interface. The available 22 full-length amino acid sequences for members of the transaldolase subfamily were aligned using ClustalW (http://www2.ebi.ac.uk/clustalw).

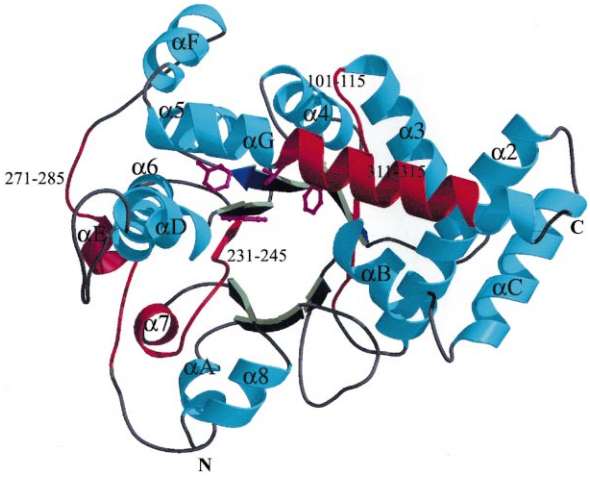

Fig. 2. Schematic view of the transaldolase monomer. The four most immunodominant peptide stretches in MS patients are shown in red and residue numbers are given. The side chains participating in the invariant hydrophobic cluster are shown in magenta. The figure was generated with Bobscript [24] and Raster3d [25].

The X-ray data were processed and merged with the programs DENZO [11] and SCALA [12]. Cell dimensions and space group (Table 1) were determined by the auto-indexing routine in DENZO and by manual inspection of simulated precession images using the program PATTERN [13]. Self-rotation functions indicated the presence of a local two-fold axis suggesting a dimer in the crystal asymmetric unit, consistent with a V_M value of 2.37 Å^3/Da.

The structure was determined by molecular replacement using the program AMORE [14] and the coordinates for transaldolase from *E. coli* [4] as search model. In a first run, the orientation and position of one of the two subunits in the asymmetric unit was found and kept fixed when searching for the second molecule. Both rotation and translation searches were calculated with data in the resolution interval 20.0–4.0 Å with an integration radius of 30 Å. The search using one subunit gave two solutions with a correlation coefficient 0.458 and an R value of 0.483. After inclusion of the second molecule these values changed to 0.614 and 0.415, respectively.

2.3. Model building and refinement

After initial rigid body refinement using the program REFMAC [15] the side chains were exchanged for their proper counterparts [9] in the human enzyme. Further refinement was performed in the interval 24–2.45 Å with the program CNS [16] followed by manual adjustments after each refinement cycle using the programs O [17] and Oops [18]. Tight non-crystallographic symmetry restraints were imposed, except for residues 11–12, 45–47, 51–53, 69–79, 123, 202–206, 230, 265, 276–285, 325 and 332. In these cases the electron density maps indicated deviations from the two-fold symmetry. In the refinement the parameters described by Engh and Huber [19] were used. The quality of the model was checked with the program PROCHECK [20]. Detailed statistics of refinement and model are listed in Table 1. The observed structure factor amplitudes and the atomic coordinates have been deposited with the Protein Data Bank, accession code 1fo5.

3. Results and discussion

3.1. Structure determination and electron density maps

In the electron density maps, the polypeptide chain is well defined except for 10 residues at the N-terminus and five residues at the C-terminus. The quality of the electron density map was very good for the A chain, while the B chain showed a few regions with weak electron density. In both chains, residue Ser-237 is found in the generously allowed region of the Ramachandran plot. This residue is well defined in electron

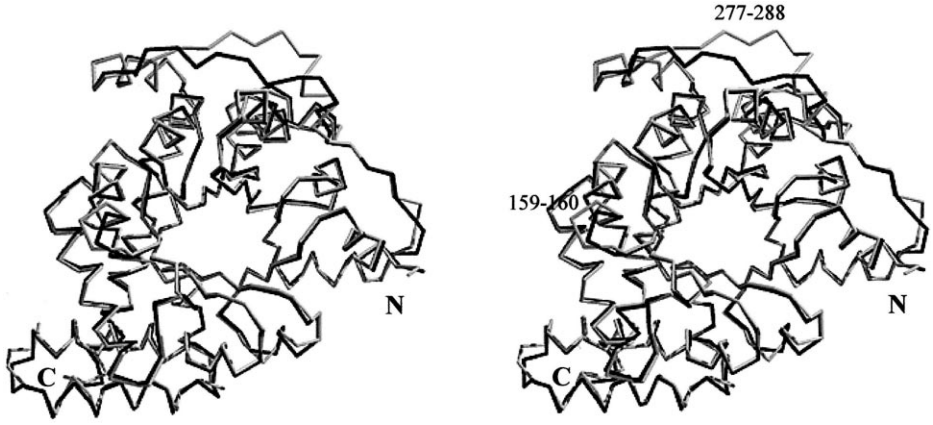

Fig. 3. Superposition of the Cα traces of human (gray) and *E. coli* (black) transaldolase. The labels indicate the peptide segments which differ most in structure between the two enzymes. The figure was generated using the programs Bobscript [24] and Raster3d [25].

Table 2
Location and suggested function of highly conserved (>90%) and invariant (bold) residues in the transaldolase subfamily

Residue	Location	Possible function
Asp-27	β1	at active site, involved in carbon–carbon bond cleavage
Thr-43	β2	close to active site, hydrogen bond to Asp-27
Asn-45	turn β2–αB	at active site, interacts with Schiff base intermediate
Pro-46	turn β2–αB	turn at active site, structurally important for proper position of Asn-45
Ala-52	αB	packing restraints, not enough space for larger side chain
Gly-101	loop α2–β3	turn at N-terminal end of barrel, structurally important
Ser-104	β3	close to active site, forms hydrogen bond to Asn-165
Glu-106	β3	at active site, involved in Schiff base formation; hydrogen bonded to water
Ala-122	α3	intramolecular hydrophobic interactions, part of conserved hydrophobic cluster (Ile-141, Ile-162)
Ile-141	β4	intramolecular hydrophobic interactions, part of hydrophobic cluster (Ala-122, Ile-162)
Lys-142	β4	at active site, Schiff base forming lysine
Thr-146	turn β4–α4	structural, hydrogen bond from side chain Oγ to main chain nitrogen of Gly-149
Gly-149	α4	space restraints after tight turn following β4
Ile-162	loop α4–β5	intramolecular hydrophobic interactions, part of conserved hydrophobic cluster (Ala-122, Ile-141)
Asn-165	β5	at active site, interacts with Schiff base intermediate
Thr-167	β5	at active site, catalytic residue
Leu-168	β5	part of invariant hydrophobic cluster
Phe-170	loop β5–α5	part of invariant hydrophobic cluster
Ala-181	loop α5–β5	packing at N-terminal end of α/β barrel
Ser-187	β6	at active site, interacts with Schiff base intermediate
Phe-189	loop β6–αD	part of invariant hydrophobic cluster
Val-190	loop β6–αD	packing of β6 against loop following α7
Arg-192	αD	at active site, part of proposed phosphate binding site
Tyr-224	α6	packing at N-terminal part of α/β barrel
Thr-231	loop α6–β7	packing at N-terminal part of α/β barrel
Leu-247	loop α7–β8	packing of β6 against loop following α7
Pro-257	α8	structurally important, beginning of helix α8 after tight turn
Gly-311	αG	part of invariant hydrophobic cluster
Phe-315	αG	part of invariant hydrophobic cluster
Leu-322	αG	packing of helix αG against αB

density in both chains and is also located in the same ϕ,ψ region in *E. coli* transaldolase.

3.2. Overall structure

Human transaldolase consists of a single domain of 337 amino acids. The core structure is an eight-stranded α/β barrel made up of eight parallel β-strands (β1–β8) and seven α-helices (α2–α8). There are seven additional α-helices (αA–αG) (Figs. 1 and 2).

In the crystal, human transaldolase forms a dimer and the interface covers an area of 1700 Å^2 corresponding to 7% of the total accessible surface area in the dimer. Contributing to the subunit–subunit interface are 18 residues from the A chain and their symmetry mates in the B chain. Most of these residues are located in helices αG, αF and the loop connecting these helices. The remaining residues are found in the loop between β3 and α3 and at the N-terminus of helix α4 (Fig. 1).

3.3. Comparison between E. coli and human transaldolase

The structures of *E. coli* and human transaldolase were compared using the lsq option in the program O with default parameters [17], giving a rms (root mean square) deviation of 0.71 Å for 309 equivalent Cα atoms. A stereo view of the superposed Cα traces is shown in Fig. 3. The peptide segments which differ most in structure in the two enzymes are residues 159–160 in the loop between α4 and β5 and residues 277–288 including αE (276–280) which is a loop in the *E. coli* enzyme. Some of the differences could be due to crystal packing, since residues 159–160 and 285–288 are engaged in intermolecular lattice contacts.

The core structure of human transaldolase differs from the *E. coli* counterpart in the number of α-helices in the α/β barrel. The α/β barrel of the human enzyme consists of eight β-strands and seven α-helices (helix α1 missing) while in the *E. coli* enzyme there is a complete barrel with eight β-strands and eight α-helices. There is also an additional helix, αE (Fig. 1), in the human enzyme, which corresponds to a loop in the *E. coli* enzyme. The subunit–subunit interactions and the dimer interface are highly conserved in the two enzymes. Out of the 18 residues found in the interface between the two subunits of *E. coli* transaldolase, nine are conserved in the human enzyme. We note, however, that none of these residues is invariant in all known sequences and only one residue, Arg-110, is more than 80% conserved in the transaldolase subfamily.

3.4. Invariant residues in the transaldolase subfamily

Amino acid sequence alignment indicates that the transaldolase family can be subdivided into two subfamilies. The 'classical' transaldolase subfamily is represented by human and *E. coli* transaldolase, and is the only subclass with a known 3D structure. Transaldolases from plants and some cyanobacteria, having large insertions/deletions at several positions, might also belong to this subclass. A more divergent subfamily, MipB/TalC, is represented by prokaryotic enzymes with a significantly shorter polypeptide chain of approximately 200 amino acids [21].

Structure-based sequence alignment of the 22 available sequences for members of the transaldolase subfamily reveals 20 invariant residues and 10 residues being at least 90% conserved (Fig. 1 and Table 2). Most of these, 27, are located in the α/β barrel core structure. These residues contribute to catalysis, substrate binding and the hydrophobic packing interactions within the α/β barrel. Eight of the invariant residues, Asp-27, Asn-45, Glu-106, Lys-142, Asn-165, Thr-167, Ser-187 and Arg-192, are involved in catalysis and substrate binding and their function is discussed in more detail elsewhere [2,4,10,22]. The invariant proline residue at position 46 is probably required to maintain the conformation of Asn-45, an invariant active site residue that interacts with the Schiff base intermediate through a hydrogen bond [2].

Of the remaining 22 invariant/highly conserved residues in the subfamily, 16 are hydrophobic amino acids which contribute to the packing of β-strands of the barrel to surrounding α-helices. Particularly noteworthy is the invariant hydrophobic cluster involving residues Leu-168, Phe-170, Phe-189, Gly-311,

and Phe-315. This small hydrophobic core is formed at the packing interface of helix αG and β-strands β5 and β6 (Fig. 2). Another small conserved hydrophobic cluster is found between helix α3, β4 and the loop close to the beginning of β5.

3.5. Human transaldolase epitopes

A systematic study identified several peptide sequences in human transaldolase that were recognized by autoantibodies from MS patients, in particular the four peptide segments comprising residues 101–115, 231–245, 271–285 and 311–325 [23]. The location of these immunodominant peptides in the 3D structure of human transaldolase is shown in Fig. 2. These segments of the polypeptide chain all contain solvent-exposed residues and are thus accessible to antibodies. Antibodies binding to the most prominent epitope, residues 271–285, showed cross-reactivity with Epstein–Barr and herpes simplex virus type 1 capsid-derived peptides suggesting that molecular mimicry might be involved in the autoimmune response [23].

Acknowledgements: We gratefully acknowledge access to synchrotron radiation at beamline B711 at the MAX Laboratory, University of Lund, Sweden. This work was supported by a grant from the Swedish Natural Science Research Council and in part by Grants RO1 DK49221 from the National Institutes of Health and RG 2466 from the National Multiple Sclerosis Society.

References

[1] Lai, C.Y., Chen, C. and Tsolas, O. (1967) Arch. Biochem. Biophys. 121, 790–797.
[2] Jia, J., Schörken, U., Lindqvist, Y., Sprenger, G.A. and Schneider, G. (1997) Protein Sci. 6, 119–124.
[3] Gefflaut, T., Blonski, C., Perie, J. and Willson, M. (1995) Prog. Biophys. Mol. Biol. 63, 301–340.
[4] Jia, J., Huang, W., Schörken, U., Sahm, H., Sprenger, G.A., Lindqvist, Y. and Schneider, G. (1996) Structure 4, 715–724.
[5] Lindqvist, Y. and Schneider, G. (1997) Curr. Opin. Struct. Biol. 7, 422–427.
[6] Banki, K., Colombo, E., Sia, F., Halladay, D., Mattson, D.H., Tatum, A.H., Massa, P.T., Philips, P.E. and Perl, A. (1994) J. Exp. Med. 180, 1649–1663.
[7] Colombo, E., Banki, K., Tatum, A.H., Daucher, J., Ferrante, P., Murray, S.M., Philips, P.E. and Perl, A. (1997) J. Clin. Invest. 99, 1238–1250.
[8] Schmidt, S. (1999) Mult. Scler. 5, 147–160.
[9] Banki, K., Halladay, D. and Perl, A. (1994) J. Biol. Chem. 269, 2847–2851.
[10] Banki, K. and Perl, A. (1996) FEBS Lett. 378, 161–165.
[11] Otwinowski, Z. (1993) DENZO: An Oscillation Data Processing Program for Macromolecular Crystallography, Yale University, New Haven, CT.
[12] Collaborative Computational Project No. 4 (1994) Acta Crystallogr. D 50, 760–763.
[13] Lu, G. (1999) J. Appl. Crystallogr. 32, 375–376.
[14] Navaza, J. (1994) Acta Crystallogr. A 50, 157–163.
[15] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Acta Crystallogr. D 53, 240–255.
[16] Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, N., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Acta Crystallogr. D 54, 905–921.
[17] Jones, T.A., Zou, J.Y., Cowan, S. and Kjeldgaard, M. (1991) Acta Crystallogr. A 47, 110–119.
[18] Kleywegt, G.J. and Jones, T.A. (1996) Acta Crystallogr. D 52, 829–832.
[19] Engh, R.A. and Huber, R. (1991) Acta Crystallogr. A 47, 392–400.
[20] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[21] Reizer, J., Reizer, A. and Saier, M.H. (1995) Microbiology 141, 961–971.
[22] Miosga, T., Schaff-Gerstenschläger, I., Franken, E. and Zimmermann, F.K. (1993) Yeast 9, 1241–1249.
[23] Esposito, M., Venkatesh, V., Otvos, L., Weng, Z., Vajda, S., Banki, K. and Perl, A. (1999) J. Immunol. 163, 4027–4032.
[24] Esnouf, R.M. (1997) J. Mol. Graph. 15, 132–134.
[25] Merritt, E.A. and Murphy, M.E.P. (1994) Acta Crystallogr. D 50, 869–873.