# = Catalytic triad =

A catalytic triad refers to the three amino acid residues that function together at the centre of the active site of some hydrolase and transferase enzymes ( e.g. proteases , amidases , esterases , acylases , lipases and ? @-@ lactamases ) . An Acid @-@ Base @-@ Nucleophile triad is a common motif for generating a nucleophilic residue for covalent catalysis . The residues form a charge @-@ relay network to polarise and activate the nucleophile , which attacks the substrate , forming a covalent intermediate which is then hydrolysed to regenerate free enzyme . The nucleophile is most commonly a serine or cysteine amino acid , but occasionally threonine . Because enzymes fold into complex three @-@ dimensional structures , the residues of a catalytic triad can be far from each other along the amino @-@ acid sequence ( primary structure ) , however , they are brought close together in the final fold .

As well as divergent evolution of function ( and even the triad 's nucleophile ) , catalytic triads show some of the best examples of convergent evolution . Chemical constraints on catalysis have led to the same catalytic solution independently evolving in at least 23 separate superfamilies . Their mechanism of action is consequently one of the best studied in biochemistry .

## = = History = =

The enzymes trypsin and chymotrypsin were first purified in the 1930s . A serine in each of trypsin and chymotrypsin was identified as the catalytic nucleophile ( by diisopropyl fluorophosphate modification ) in the 1950s . The structure of chymotrypsin was solved by X @-@ ray crystallography in the 1960s , showing the orientation of the catalytic triad in the active site . Other proteases were sequenced and aligned to reveal a family of related proteases , now called the S1 family . Simultaneously , the structures of the evolutionarily unrelated papain and subtilisin proteases were found to contain analogous triads . The ' charge @-@ relay ' mechanism for the activation of the nucleophile by the other triad members was proposed in the late 1960s . As more protease structures were solved by X @-@ ray crystallography in the 1970s and 80s , homologous ( such as TEV protease ) and analogous ( such as papain ) triads were found . The MEROPS classification system in the 1990s and 2000s began classing proteases into structurally related enzyme superfamilies and so acts as a database of the convergent evolution of triads in over 20 superfamilies . Understanding how chemical constraints on evolution led to the convergence of so many enzyme families on the same triad geometries has developed in the 2010s . Of particular contention during the 1990s and early 2000s was the contribution of low @-@ barrier hydrogen bonding to catalysis ; however , current thinking is that ordinary hydrogen bonding is sufficient to explain the mechanism . The massive body of work on the charge @-@ relay , covalent catalysis used by catalytic triads has led to the mechanism being the best characterised in all of biochemistry .

## = = Function = =

Enzymes that contain an catalytic triad use it for one of two reaction types : either to split a substrate ( hydrolases ) or to transfer one portion of a substrate over to a second substrate ( transferases ) . Triads are an inter @-@ dependent set of residues in the active site of an enzyme and act in concert with other residues ( e.g. binding site and oxyanion hole ) to achieve nucleophilic catalysis . These triad residues act together to make the nucleophile member highly reactive , generating a covalent intermediate with the substrate that is then resolved to complete catalysis .

### = = = Mechanism = = =

Catalytic triads perform covalent catalysis using a residue as a nucleophile . The reactivity of the nucleophilic residue is increased by the functional groups of the other triad members . The nucleophile is polarised and oriented by the base , which is itself bound and stabilised by the acid .

Catalysis is performed in two stages . First , the activated nucleophile attacks the carbonyl carbon and forces the carbonyl oxygen to accept an electron , leading to a tetrahedral intermediate . The build @-@ up of negative charge on this intermediate is typically stabilized by an oxanion hole within the active site . The intermediate then collapses back to a carbonyl , ejecting the first half of the substrate , but leaving the second half still covalently bound to the enzyme as an acyl @-@ enzyme intermediate . The ejection of this first leaving group is often aided by donation of a proton by the base .

The second stage of catalysis is the resolution of the acyl @-@ enzyme intermediate by the attack of a second substrate . If this substrate is water then the result is hydrolysis ; if it is an organic molecule then the result is transfer of that molecule onto the first substrate . Attack by this second substrate forms a new tetrahedral intermediate , which resolves by ejecting the enzyme 's nucleophile , releasing the second product and regenerating free enzyme .

= = Identity of triad members = =


= = = Nucleophile = = =

The side @-@ chain of the nucleophilic residue performs covalent catalysis on the substrate . The lone pair of electrons present on the oxygen or sulphur attacks the electropositive carbonyl carbon . The 20 naturally occurring biological amino acids do not contain any sufficiently nucleophilic functional groups for many difficult catalytic reactions . Embedding the nucleophile in a triad increases its reactivity for efficient catalysis . The most commonly used nucleophiles are the hydroxyl ( OH ) of serine and the thiol / thiolate ion ( SH / S ? ) of cysteine . Alternatively , threonine proteases use the secondary hydroxyl of threonine , however due to steric hindrance of the side chain 's extra methyl group such proteases use their N @-@ terminal amide as the base , rather than a separate amino acid .

Use of oxygen or sulphur as the nucleophilic atom causes minor differences in catalysis . Compared to oxygen , sulphur ? s extra d orbital makes it larger ( by 0 @.@ 4 Å ) and softer , allows it to form longer bonds ( dC @-@ X and dX @-@ H by 1 @.@ 3 @-@ fold ) , and gives it a lower pKa ( by 5 units ) . Serine is therefore more dependent than cysteine on optimal orientation of the acid @-@ base triad members to reduce its pKa in order to achieve concerted deprotonation with catalysis . The low pKa of cysteine works to its disadvantage in the resolution of the first tetrahedral intermediate as unproductive reversal of the original nucleophilic attack is the more favourable breakdown product . The triad base is therefore preferentially oriented to protonate the leaving group amide to ensure that it is ejected to leave the enzyme sulphur covalently bound to the substrate N @-@ terminus . Finally , resolution of the acyl @-@ enzyme ( to release the substrate C @-@ terminus ) requires serine to be re @-@ protonated whereas cysteine can leave as S ? . Sterically , the sulphur of cysteine also forms longer bonds and has a bulkier Van der Waals radius and if mutated to serine can be trapped in unproductive orientations in the active site .

= = = Base = = =

Since no natural amino acids are strongly nucleophilic , the base in a catalytic triad polarises and deprotonates the nucleophile to increase its reactivity . Additionally , it protonates the first product to aid leaving group departure .

The base is most commonly histidine since its pKa allows for effective base catalysis , hydrogen bonding to the acid residue , and deprotonation of the nucleophile residue. ? @-@ lactamases such as TEM @-@ 1 use a lysine residue as the base . Because lysine 's pKa is so high ( pKa = 11 ) , a glutamate and several other residues act as the acid to stabilise its deprotonated state during the catalytic cycle . Threonine proteases use their N @-@ terminal amide as the base , since steric crowding by the catalytic threonine 's methyl prevents other residues from being close enough .

### = = = Acid = = =

The acidic triad member forms a hydrogen bond with the basic residue . This aligns the basic residue by restricting its side @-@ chain rotation , and polarises it by stabilising its positive charge . Two amino acids have acidic side chains at physiological pH ( aspartate or glutamate ) and so are the most commonly used for this triad member . Cytomegalovirus protease uses a pair of histidines , one as the base , as usual , and one as the acid . The second histidine is not as effective an acid as the more common aspartate or glutamate , leading to a lower catalytic efficiency . In some enzymes , the acid member of the triad is less necessary and some act only as a dyad . For example , papain uses asparagine as its third triad member which orients the histidine base but does not act as an acid . Similarly , hepatitis A virus protease contains an ordered water in the position where an acid residue should be .

## = = Examples of triads = =

### = = = Ser @-@ His @-@ Asp = = =

The Serine @-@ Histidine @-@ Aspartate motif is one of the most thoroughly characterised catalytic motifs in biochemistry . The triad is exemplified by chymotrypsin , a model serine protease from the PA superfamily which uses its triad to hydrolyse protein backbones . The aspartate is hydrogen bonded to the histidine , increasing the pKa of its imidazole nitrogen from 7 to around 12 . This allows the histidine to act as a powerful general base and to activate the serine nucleophile . It also has an oxanion hole consisting of several backbone amides which stabilises charge build @-@ up on intermediates . The histidine base aids the first leaving group by donating a proton , and also activates the hydrolytic water substrate by abstracting a proton as the remaining OH ? attacks the acyl @-@ enzyme intermediate .
The same triad has also convergently evolved in ? / ? hydrolases such as some lipases and esterases , however orientation of the triad members is reversed . Additionally , brain acetyl hydrolase ( which has the same fold as a small G @-@ protein ) has also been found to have this triad . The equivalent Serine @-@ Histidine @-@ Glutamate triad is used in acetylcholinesterase .

### = = = Cys @-@ His @-@ Asp = = =

The second most studied triad is the Cysteine @-@ Histidine @-@ Aspartate motif . Several families of cysteine proteases use this triad set , for example TEV protease and papain . The triad acts similarly to serine protease triads , with a few notable differences . Due to cysteine 's low pKa , the importance of the Asp to catalysis varies and several cysteine proteases are effectively Cys @-@ His dyads ( e.g. hepatitis A virus protease ) , whilst in others the cysteine is already deprotonated before catalysis begins ( e.g. papain ) . This triad is also used by some amidases , such as N @-@ glycanase to hydrolyse non @-@ peptide C @-@ N bonds .

### = = = Ser @-@ His @-@ His = = =

The triad of cytomegalovirus protease uses histidine as both the acid and base triad members . Removing the acid histidine results in only a 10 @-@ fold activity loss ( compared to > 10 @,@ 000 @-@ fold when aspartate is removed from chymotrypsin ) . This triad has been interpreted as a possible way of generating a less active enzyme to control cleavage rate .

### = = = Ser @-@ Glu @-@ Asp = = =

An unusual triad is found in seldolisin proteases . The low pKa of the glutamate carboxylate group means that it only acts as a base in the triad at very low pH . The triad is hypothesised to be an

adaptation to specific environments like acidic hot springs ( e.g. kumamolysin ) or cell lysosome ( e.g. tripeptidyl peptidase ) .

### = = = Thr @-@ Nter , Ser @-@ Nter and Cys @-@ Nter = = =

 Threonine proteases , such as the proteasome protease subunit and ornithine acyltransferases use the secondary hydroxyl of threonine in an manner analogous to the use of the serine primary hydroxyl . However , due to the steric interference of the extra methyl group of threonine , the base member of the triad is the N @-@ terminal amide which polarises an ordered water which , in turn , deprotonates the catalytic hydroxyl to increase its reactivity . Similarly , there exist equivalent ' serine only ' and ' cysteine only ' configurations such as penicillin acylase G and penicillin acylase V which are evolutionarily related to the proteasome proteases . Again , these use their N @-@ terminal amide as a base .

### = = = Ser @-@ cisSer @-@ Lys = = =

 This unusual triad occurs only in one superfamily of amidases . In this case , the lysine acts to polarise the middle serine . The middle serine then forms two strong hydrogen bonds to the nucleophilic serine to activate it ( one with the side chain hydroxyl and the other with the backbone amide ) . The middle serine is held in an unusual cis orientation to facilitate precise contacts with the other two triad residues . The triad is further unusual in that the lysine and cis @-@ serine both act as the base in activating the catalytic serine , but the same lysine also performs the role of the acid member as well as making key structural contacts .

## = = Divergent evolution = =

 The sophistication of the active site network causes residues involved in catalysis ( and residues in contact with these ) to be highly evolutionarily conserved . However , there are examples of divergent evolution in catalytic triads , both in the reaction catalysed , and the residues used in catalysis . The triad remains the core of the active site , but it is evolutionarily adapted to serve different functions .

### = = = Reaction changes = = =

 Catalytic triads perform covalent catalysis via an acyl @-@ enzyme intermediate . If this intermediate is resolved by water , the result is hydrolysis of the substrate . However , if the intermediate is resolved by attack by a second substrate , then the enzyme acts as a transferase . For example , attack by an acyl group results in an acyltransferase reaction . Several families of transferase enzymes have evolved from hydrolases by adaptation to exclude water and favour attack of a second substrate .
 Additionally , an alternative transferase mechanism has been evolved by amidophosphoribosyltransferases , which has two active sites . In the first active site , a cysteine triad hydrolyses a glutamine substrate to release free ammonia . The ammonia then diffuses though an internal tunnel in the enzyme to the second active site , where it is transferred to a second substrate .

### = = = Nucleophile changes = = =

 Divergent evolution of active site residues is slow , due to strong chemical constraints . Nevertheless , some protease superfamilies have evolved from one nucleophile to another . This can be inferred when a superfamily ( with the same fold ) contains families that use different nucleophiles . Such nucleophile switches have occurred several times during evolutionary history , however the mechanisms by which this happen are still unclear .

Within protease superfamilies that contain a mixture of nucleophiles ( e.g. the PA clan ) , families are designated by their catalytic nucleophile ( C = cysteine proteases , S = serine proteases ) .

## = = Convergent evolution = =

The enzymology of proteases provides some of the clearest known examples of convergent evolution . The same geometric arrangement of triad residues occurs in over 20 separate enzyme superfamilies . Each of these superfamilies is the result of convergent evolution for the same triad arrangement within a different structural fold . This is because there are limited productive ways to arrange three triad residues , the enzyme backbone and the substrate . These examples reflect the intrinsic chemical and physical constraints on enzymes , leading evolution to repeatedly and independently converge on equivalent solutions .

### = = = Cysteine and serine hydrolases = = =

The same triad geometries been converged upon by serine proteases such as the chymotrypsin and subtilisin superfamilies . Similar convergent evolution has occurred with cysteine proteases such as viral C3 protease and papain superfamilies . These triads have converged to almost the same arrangement due to the mechanistic similarities in cysteine and serine proteolysis mechanisms .
Families of Cysteine proteases
Families of Serine proteases

### = = = Threonine proteases = = =

Threonine proteases use the amino acid threonine as their catalytic nucleophile . Unlike cysteine and serine , threonine is a secondary hydroxyl ( i.e. has a methyl group ) . This methyl group greatly restricts the possible orientations of triad and substrate as the methyl clashes with either the enzyme backbone or histidine base . Consequently , the catalytic residue of a threonine protease is located at it N @-@ terminus .
Two evolutionarily independent enzyme superfamilies with different protein folds are known to use the N @-@ terminal residue as a nucleophile : Superfamily PB ( proteasomes using the Ntn fold ) and Superfamily PE ( acetyltransferases using the DOM fold ) This commonality of active site structure in completely different protein folds indicates that the active site evolved convergently in those superfamilies .
Families of threonine proteases